Radiation from collapsing shells, semiclassical backreaction and black hole formation

Aseem Paranjape*
Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai - 400 005, India

T. Padmanabhan†
IUCAA, Post Bag 4, Ganeshkhind, Pune - 411 007, India

(Dated: July 23, 2009)

We provide a detailed analysis of quantum field theory around a collapsing shell and discuss several conceptual issues related to the emission of radiation flux and formation of black holes. Explicit calculations are performed using a model for a collapsing shell which turns out to be analytically solvable. We use the insights gained in this model to draw reliable conclusions regarding more realistic models. We first show that any shell of mass M which collapses to a radius close to $r = 2M$ will emit approximately thermal radiation for a period of time. In particular, a shell which collapses from some initial radius to a final radius $2M(1-\epsilon^2)^{-1}$ (where $\epsilon \ll 1$) without forming a black hole, will emit thermal radiation during the period $M \lesssim t \lesssim M\ln(1/\epsilon^2)$. Later on ($t \gg M\ln(1/\epsilon^2)$), the flux from such a shell will decay to zero exponentially. We next study the effect of backreaction computed using the vacuum expectation value of the stress tensor on the collapse. We find that, in any realistic collapse scenario, the backreaction effects do *not* prevent the formation of the event horizon. The time at which the event horizon is formed is, of course, delayed due to the radiated flux — which decreases the mass of the shell — but this effect is not sufficient to prevent horizon formation. We also clarify several conceptual issues and provide pedagogical details of the calculations in the Appendices to the paper.

I. INTRODUCTION, MOTIVATION AND SUMMARY

Classical general relativity allows for solutions in which matter can collapse in a spherically symmetric manner, forming a black hole event horizon when viewed from the outside region. The collapsing matter hits a spacetime singularity in finite proper time thereby preventing us from using the classical equations to predict the future evolution of the system, as viewed by an observer collapsing with the matter. On the other hand, as far as any outside observer is concerned, the collapsing material takes infinite time (as measured by the stationary clocks outside) to reach the event horizon and hence the formation of the singularity has no influence on the outside region. While this may appear bizarre to the uninitiated, general relativists have learned to live with this dichotomous evolutionary behaviour of the system.

A new layer of complication arises when we go beyond the classical theory and study quantum fields evolving in the background geometry of collapsing matter. If a quantum field is in the vacuum state in the asymptotic past, then at late times an observer at spatial infinity will see a flux of energy from the collapsing matter corresponding to a thermal radiation with temperature $T_H = 1/8\pi M$ [1]. (Throughout the paper, we will use Planckian units with G, $\hbar$ and c set to unity.) The thermal nature is closely related to the exponential redshift experienced by the modes when they travel from a region close to $r = 2M$ to $r \to \infty$ and — to that extent — the thermal nature of the radiation depends on matter collapsing to a size $r \approx 2M$. This process has been extensively investigated and is well understood as long as we treat the quatum field as a test field (see Ref. [2] for a pedagogical treatment and a review of the literature).

But the radiation flux to $r \to \infty$ is 'real' in the sense that the observer can collect it and use it to do work (say, e.g., to heat up some water). Energy conservation then requires that the mass of the collapsing body must decrease due to the outgoing radiation, since it is the only possible source of energy. In that case, we cannot treat the background as fixed but must solve for Einstein's equations with a semiclassical backreaction term added (which is usually modeled through the vacuum expectation value (VEV) of the stress tensor with the adjective 'vacuum' referring to the choice that, in the asymptotic past, the quantum field was in the vacuum state) and solve for the radiated flux simultaneously and self-consistently. Several authors have studied this phenomenon [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16] and

*Electronic address: aseem@tifr.res.in
†Electronic address: paddy@iucaa.ernet.in

in particular we refer to the review by Brout et al. [3]. As long as the flux of particles, characterized by the VEV of the stress tensor is *small* and *slowly varying* (in Planckian units), the backreaction can be self-consistently accounted for and studied in the geometry of a black hole with a *slowly varying mass.* In such a case, the backreaction effects are not significant and can be incorporated in a perturbative matter.

Recently there have been claims in the literature [17] that the number of particles reaching an observer at large distances from the black hole, at late times, can diverge. It has then been conjectured that the backreaction from this diverging number of particles might be sufficient to prevent the formation of the event horizon. These claims are clearly contradictory to the usually accepted view mentioned above, and other authors have also questioned this view (see, e.g. Refs. [18, 19]; see also Ref. [20]). In the light of such claims, we wish to revisit this issue in this paper.

Right at the outset, let us clarify one elementary — but potentially misleading — aspect of the problem. During the classical gravitational collapse, although the collapsing object will cross its Schwarzschild radius (and thus form an event horizon) in finite proper time, an observer at rest at large distances from the collapsing object will see the collapsing body approach its Schwarzschild radius only asymptotically as $R(t) = 2M(1 + \mathcal{O}(e^{-t/4M}))$ where t is the usual Schwarzschild time coordinate of the outside metric

$$ds^2_{\rm Schw} = -(1 - 2M/r)dt^2 + \frac{dr^2}{(1 - 2M/r)} + r^2 d\Omega^2 \,. \tag{1}$$

(Throughout the paper, we shall deal with situations in which a spherical body collapses without rotation.) In the absence of any backreaction, a calculation of the stress tensor VEV will show a steady flux at late times corresponding to a (nearly) thermal spectrum with a temperature $T_H \propto 1/M$. If one insists on waiting for an arbitrarily long time *without accounting for backreaction*, we will clearly end up with an arbitrarily large number of particles reaching large distances at late enough times. Such a divergence, of course, has nothing to do with black holes, and will arise for *any* luminous object if we pretend that its mass does not reduce due to the energy it radiates. For example, if the sun could radiate a steady flux of photons for an infinite amount of time without changing its mass or constitution, it would emit an infinite amount of energy! That is, this situation will arise regardless of whether the event horizon forms or not and the divergence merely indicates the need to correctly account for the backreaction (or, more simply, energy conservation).

In principle, it is therefore inconsistent to assume that the exterior geometry is described by an exact Schwarzschild metric for two reasons: First, the mass of the central object is changing and second, the presence of an outgoing flux even at large distances means that the geometry is not strictly asymptotically flat. One could now imagine a situation like the following : the mass loss of the collapsing object (in our case, the shell) implies that the radius to which the shell must collapse to form an event horizon, is shrinking. This means that the shell spends a larger amount of time outside the event horizon, all the while radiating particles to infinity. All this is likely to be further complicated by the fact that if the mass loss rate is high enough, the exterior geometry is likely to be very complicated. The radius which the shell is chasing keeps reducing, and the possibility arises that this runaway process ends with the shell completely evaporating *before* the event horizon is formed. If this situation is generic, it would mean that black holes typically *do not* form, contrary to the popular view. It is therefore worth investigating this scenario to settle the effect of backreaction, and in this paper we will do this using some simple models for collapse. The simplifications we introduce to make the problem tractable are the following:

(a) We assume that the collapsing system is a thin spherical shell with internal stresses arranged in such a manner as to allow it to collapse along some given trajectory in the spacetime. Then the internal metric is Minkowski and the outside is Schwarszchild with matching conditions determining either metric in terms of the other and the trajectory of the shell [11].

(b) We will assume that all computations can be performed in the $1 + 1$ sector of the spacetime which ignores the two angular coordinates. This corresponds to using only the s-wave mode of the scalar field and allows one to use the tools of conformal field theory to compute the VEV of the stress tensor. This assumption is unlikely to influence the conclusions of the paper and is justified in Ref. [3].

(c) We will assume that the semiclassical backreaction can be modeled by the VEV of the stress tensor, $< T_{ab} >$ with Einstein's equations modified to $G_{ab} = 8\pi(T_{ab}+ < T_{ab} >)$. There is general agreement that such an equation should be valid in some suitable limit (though no one has rigorously proved it; but see Ref. [21]) and we will proceed hoping for the best.

Given these assumptions, the problem can in principle be reduced to one of solving a set of equations. Given any trajectory of the shell, one can compute the VEV of T_{ab} everywhere using the conformal field theory technique. The problem then reduces to calculating (a) the backreaction on the shell trajectory and (b) the flux radiated to infinity. In practice, unfortunately, the equations turn out to be quite intractable and one needs to obtain insights into what is happening using simplified models. We do this along the following lines in this paper.

We first consider the role of the event horizon in the emission of particles to large distances (ignoring the effect of the backreaction). We conclude that operationally, its role depends on when the observer at infinity sets out to detect the particles. To understand this result, consider two collapse scenarios (say, A and B) in which shells (each

of mass M) start from the same initial radius R_0 with further evolution being different: (a) In Case A, the shell continuously collapses and forms an event horizon. (The shell crosses $r = 2M$ at a finite proper time as shown by the clock on the surface of the shell but after infinite amount of coordinate time as measured by the clocks of the observer outside the shell.) Standard calculations indicate that at late times the asymptotic observer will see a thermal flux of radiation. (b) In Case B, the shell follows exactly the same trajectory as in Case A until it reaches close to $r = 2M$. Its trajectory then deviates from that of Case A; the shell progressively slows down, and asymptotically approaches a final radius $r = 2M(1 - \epsilon^2)^{-1}$. Obviously no event horizon is formed in this Case B for any value of ϵ, however small. It is also clear that at very late times the geometry is static in Case B with the shell staying at a fixed radius outside the event horizon and there should be no flux of radiation. But for an arbitrarily large period of time we can arrange matters such that the trajectory in Case B behaves similar to that in Case A when it is hovering just outside $r = 2M$ (as seen by an outside observer). The outgoing modes which cross the shell and propagate to future null infinity will lead to a thermal flux in this case as well. In fact, this is an elementary consequence of causality. What an observer at infinity sees at some given event $\mathcal{P}$ (at time t_{obs}) can only depend on the behaviour of shell trajectory which is contained within the backward light cone of $\mathcal{P}$. The future trajectory of the shell — in particular whether it settles down at $r \gtrsim 2M$ or goes on to collapse through $r = 2M$ — should not affect observations at $\mathcal{P}$.

We therefore expect the Case B to be characterized by three distinct phases. In the first phase, the shell collapses from a large radius to some radius close to $2M$ but larger than its final asymptotic radius. In this phase, the trajectory of the shell is identical to that of Case A. We, therefore, expect a flux of radiation being emitted to infinity which starts from zero and builds up to the thermal flux value characterized by the temperature $T_H \propto (1/M)$ in the timescale of the collapse (which is $\mathcal{O}(M)$ for nearly geodesic motion). In the second phase, the shell hovers close to $r = 2M$ *before* settling down to its asymptotic radius. If we choose the trajectory such that the timescale governing this phase is $\mathcal{O}(M)$, then during this phase also the trajectory is similar to that of Case A, and now we would expect an approximately thermal flux being emitted by the shell just as though it is going to collapse to $r = 2M$. The crucial difference between the two cases arise in the third phase, during which the shell is asymptotically coming to rest. *We will show by explicit calculation and a numerical analysis that during this phase the flux of radiation from the body dies down to zero exponentially.* If the first two phases are characterized by a timescale $\mathcal{O}(M)$, then the intermediate phase during which the shell emits approximately thermal radiation lasts for a timescale of the order of $M \ln(1/\epsilon^2)$.

This explicit demonstration is one of the key results of this paper. This result shows that the thermal flux of radiation observed from a collapsing structure is not directly dependent on the formation of the event horizon. When the event horizon does form, the thermal radiation continues to escape for infinite amount of time if the backreaction effects are ignored. This is, of course, unrealistic since it would require an infinite source of energy. So in any realistic collapse scenario, the emission of thermal radiation is an idealization which ceases to be valid at sufficiently late times. Our model calculation shows that a similar effect can be mimicked by having a system collapse to a radius $r = 2M(1 - \epsilon^2)^{-1}$ asymptotically. The approximately thermal radiation will emanate from the body during the time interval $M \lesssim t \lesssim M \ln(1/\epsilon^2)$. At very late times ($t \gg M \ln(1/\epsilon^2)$) the radiation dies down exponentially.

The second key result of the paper is related to the computation of semiclassical backreaction and its effect on the formation of event horizon. We show that the effect of backreaction is essentially to delay the formation of event horizon by an amount δv (where v is the in-going Eddington-Finkelstein null coordinate) given by

$$\delta v = \mathcal{O}\left(\frac{M}{\beta_h^2} L_H\right) , \tag{2}$$

where β_h is the proper velocity with which the unperturbed trajectory crosses $r = 2M$ and L_H is the luminosity of the thermal radiation. This result is intuitively clear : The delay in the event horizon formation is governed by the product of L_H and the timescale of the unperturbed collapse, which is essentially set by the initial mass of the collapsing object. Physically this is simply the naive bound one might place on the amount of mass that the collapsing object can radiate via semiclassical emission, during the collapse. The velocity at (unperturbed) horizon crossing β_h is completely determined by the initial conditions (namely that the shell starts collapsing with zero velocity at a radius $r = R_0$ at initial time). For example, for geodesic infall starting at $r = R_0$, one has $\beta_h^2 = 1 - 2M/R_0$, and typically one expects $\beta_h \lesssim \mathcal{O}(1)$ for more general trajectories as well. For all such trajectories the backreaction, which varies on a timescale $\sim M^3$, *cannot prevent* the formation of the event horizon which occurs on a timescale $\sim M$; at best it can delay this formation by a time $\sim 1/M$. (One may wonder whether one can arrange matters for β_h to become sufficiently small to increase this timescale arbitrarily. This looks implausible and we provide detailed arguments in the relevant section eliminating this possibility.)

It is possible to use these results to also put a bound on the total amount of mass that can be radiated away before the event horizon is formed and we find that it is given by $M^{-1} \ln M$ for most of the realistic trajectories. These results strongly suggest that the conventional wisdom as regards the formation of black hole is correct and that the semiclassical radiation does not prevent the formation of the event horizon.

The plan of the paper is as follows. Section II deals with the computation of the renormalized stress tensor VEV for a collapsing shell, in the *absence* of backreaction. We first recall some basic concepts of quantum fields in curved

spaces in Sec. II A. To set the stage, in Sec. II B we analyse a trajectory which does not form a horizon and whose collapsing phase is null. While being unrealistic, this toy example admits of an exact calculation of the stress tensor VEV. In Sec. II C we turn to the more realistic timelike trajectory discussed above, and analyse the behaviour of the stress tensor VEV in various phases of the collapse (which, again, does not form a horizon).

In section III we include the effects of the backreaction of the stress tensor VEV, on the background geometry. We begin by recalling details of a self-consistent picture as presented by Brout et al. [3], which describes the exterior geometry in the presence of a *small* and *slowly varying* emission from the collapsing object. We then present a simple calculation based on this picture, to derive the expression (2) for the delay in formation of the event horizon in the presence of backreaction. We argue that even for (classical) trajectories which slow down drastically as they approach the unperturbed Schwarzschild radius, it is highly implausible that backreaction can significantly delay the formation of the event horizon. Finally, we calculate an upper bound on the mass that can be radiated away due to semiclassical emission, using the extreme trajectories studied in Sec. II. We briefly highlight our main results and conclude in section IV. Appendix A contains pedagogical details of results from 2-dimensional conformal field theory which are relevant to our calculations, and Appendix B contains detailed proofs of various results that are quoted in the main text.

II. FLUX OF RADIATION FROM A COLLAPSING SHELL

We begin by recalling several results related to the quantization of a scalar field propagating on a fixed background geometry, which is taken to be that of a collapsing object. At the zeroth order we treat the scalar field as a test field so that the stress-energy tensor of the scalar field does not affect the background. For simplicity, throughout the paper we will consider the collapse of a thin shell of mass M, so that the exterior geometry is described by the Schwarzschild metric, while the interior is flat Minkowski spacetime. Since we will eventually be mainly interested in order of magnitude estimates and asymptotic behaviour of various quantities, we will focus on broad properties of the collapse trajectory without going into details of the shell stress-energy tensor, etc.

A. Background geometry and mode analysis

The exterior geometry of the shell is given by the Schwarzschild metric, which we write using the Eddington-Finkelstein coordinates, as

$$\begin{aligned} ds^2_{(\text{ext})} &= -(1-\frac{2M}{r})dudv + r^2(u,v)d\Omega^2 \\ &= -(1-2M/r)dv^2 + 2dvdr + r^2 d\Omega^2 \,, \end{aligned} \tag{3}$$

where M is the (constant) mass of the shell and r is defined implicitly via

$$e^{(v-u)/4M} = \left(\frac{r}{2M} - 1\right) e^{r/2M} \quad ; \quad r > R_s \,, \tag{4}$$

where $r = R_s(\tau)$ is the trajectory of the shell with proper time τ. The interior geometry is that of Minkowski spacetime, whose metric we write as

$$ds^2_{(\text{int})} = -dT^2 + dr^2 + r^2 d\Omega^2 = -dUdV + r^2(U,V)d\Omega^2 = -dV^2 + 2dVdr + r^2 d\Omega^2 \,, \tag{5}$$

where we have defined the light cone coordinates $V = T + r$, $U = T - r$, and we use the same r coordinate to label 2-spheres in the interior and exterior. These metrics can be matched at the surface of the shell to get equations for $U(u)$ and $V(v)$ which are valid throughout the spacetime outside $r = 2M$. We will always define the shell's trajectory by the value of its r-coordinate as $r = R_s$, however we will treat R_s as a function of either u or v depending on convenience. The matching conditions are

$$\begin{aligned} &dV - dU = 2dR_s \quad ; \quad dv - du = \frac{2dR_s}{(1-2M/R_s)} \,, \\ &dUdV = \left(1 - \frac{2M}{R_s}\right) dudv \,, \end{aligned} \tag{6}$$

with the differentials understood to be along the trajectory. For a timelike trajectory, these can be used to write the equation of the trajectory in the following alternative forms which will be useful later

$$2R_s'(1-U') = U'^2 - \left(1 - \frac{2M}{R_s}\right) \,, \tag{7}$$

$$2\dot{R}_s(1-\dot{V}) = \left(1-\frac{2M}{R_s}\right) - \dot{V}^2\,, \tag{8}$$

where the prime and dot denote derivatives with respect to u and v respectively, along the trajectory. Once the trajectory is specified, these equations completely determine the metric both inside and outside and also allow us to relate the coordinates in the two regions.

We next consider the quantization of the scalar field in this spherical geometry. Rigorously speaking, one needs to introduce the spherical harmonics $Y_{lm}(\theta,\phi)$ and separate out the angular dependence of the modes. The resulting expressions will lead to a VEV of the stress tensor which cannot be evaluated analytically. To make progress, we shall concentrate only on the s-wave component of the scalar field and reduce the problem to one in the $r-t$ plane. In such a 2-dimensional context, one can use the techniques of conformal field theory to evaluate the VEV of stress-tensor. We further ignore the effects of the Schwarzschild potential barrier which arises in the exterior even for the s-wave component. (See Refs. [3, 5, 6] for a detailed discussion of the accuracy of these approximations, and Ref. [2] for elementary arguments supporting the s-wave approximation.) The scalar field is therefore taken to satisfy the 2-dimensional Klein-Gordon equation both inside and outside the shell, i.e.,

$$\begin{aligned} \partial_U\partial_V\varphi &= 0\,, \quad \text{(interior)}\,, \\ \partial_u\partial_v\varphi &= 0\,, \quad \text{(exterior)}\,. \end{aligned} \tag{9}$$

We shall now briefly recall the procedure for quantizing the scalar field in the 1+1 spacetime and collect together the relevant formulas. The general solution of the Klein-Gordon equation in the exterior region is

$$\varphi(u,v) = f(v) + \xi(u)\,, \tag{10}$$

which after matching at the shell becomes

$$\varphi(U,V) = f(v(V)) + \xi(u(U)) \equiv \tilde{f}(V) + \tilde{\xi}(U)\,, \tag{11}$$

in the interior region. Imposing the reflection condition that the solution vanish at the center $r=(V-U)/2=0$ then gives

$$\varphi = \tilde{f}(V) - \tilde{f}(U) = f(v(V)) - f(v(U))\,, \tag{12}$$

in the interior, where $v(U)$ is the function $v(V)$ evaluated at $V=U$. We think of these solutions as wave-packets and study each frequency mode separately. Denote by $\varphi_\omega^{\rm in}$ modes which are positive frequency with respect to Schwarzschild time on past null infinity $\mathcal{I}^-$. These modes will then define a vacuum (see below) which corresponds to the Minkowski vacuum on $\mathcal{I}^-$. The relevant in-falling part of $\varphi_\omega^{\rm in}$ is

$$\varphi_\omega^{\rm in} \sim \frac{1}{\sqrt{4\pi\omega}} e^{-i\omega v}\,, \quad (\text{on } \mathcal{I}^-)\,, \tag{13}$$

and is normalized with respect to the Klein-Gordon norm

$$(f_1,f_2)_{(v)} = \int_{-\infty}^{\infty} dv f_2^*(i\overleftrightarrow{\partial_v})f_1\,, \tag{14}$$

and similarly with u. The previous discussion on matching and reflection conditions shows that the outgoing u-dependent part of these modes, which reaches future null infinity $\mathcal{I}^+$, is given by

$$\varphi_\omega^{\rm in} \sim -\frac{1}{\sqrt{4\pi\omega}} e^{-i\omega G(u)}\,, \quad (\text{on } \mathcal{I}^+)\,, \tag{15}$$

where the function $G(u)$ is defined by the following chain : An ingoing mode labelled by $v=v_{\rm ent}$ behaves like $\sim e^{-i\omega v_{\rm ent}}$ in the exterior. It enters the shell at the event $\mathcal{P}_{\rm ent}$ and now behaves like $\sim e^{-i\omega v_{\rm ent}(V)}$ in the interior. The function $v_{\rm ent}(V)$ is determined by the matching conditions at $\mathcal{P}_{\rm ent}$. At reflection the mode becomes $\sim e^{-i\omega v_{\rm ent}(U)}$, and after exiting the shell at the event $\mathcal{P}_{\rm exit}$ it becomes $\sim e^{-i\omega v_{\rm ent}(U(u))}$. The function $G(u)$ is hence given by

$$G(u) = v_{\rm ent}(U(u))\,. \tag{16}$$

The scalar field is quantized by defining a vacuum state $|\texttt{in}\rangle$ with respect to these modes by writing the field operator as

$$\varphi = \int_0^\infty d\omega\,\left(a_\omega\varphi_\omega^{\rm in} + \text{h.c.}\right)\,, \tag{17}$$

where "h.c." stands for Hermitian conjugate, and the a_ω are annihilation operators for the state $|\mathtt{in}\rangle$,

$$a_\omega|\mathtt{in}\rangle = 0\,. \tag{18}$$

We will work throughout in the Heisenberg picture, so that the state does not evolve, and hence all expectation values must be computed in the state $|\mathtt{in}\rangle$.

A key point is that on $\mathcal{I}^+$, the field operator has modes which are positive frequency with respect to $G(u)$, not u. So the state $|\mathtt{in}\rangle$ is the vacuum of the G-modes, or the "G-vacuum". Since $G(u)$ is a nonlinear function of u in general, this G-vacuum will contain particles corresponding to the "u-vacuum", defined as the state annihilated by modes which are positive frequency with respect to u. We will denote this vacuum by the state $|\mathtt{out}\rangle$, annihilated by operators b_λ corresponding to modes $\varphi^{\rm out}_\lambda$, such that on $\mathcal{I}^+$, the relevant outgoing part of $\varphi^{\rm out}_\lambda$ is

$$\varphi^{\rm out}_\lambda \sim \frac{1}{\sqrt{4\pi\lambda}}e^{-i\lambda u}\,, \quad (\text{on } \mathcal{I}^+)\,, \tag{19}$$

and it should be clear that the corresponding in-falling part on $\mathcal{I}^-$ is

$$\varphi^{\rm out}_\lambda \sim -\frac{1}{\sqrt{4\pi\lambda}}e^{-i\lambda h(v)}\,, \quad \text{on } \mathcal{I}^-\,, \tag{20}$$

where if $u = h(v)$ then $v = G(u)$. The field operator φ can be expanded in the $\varphi^{\rm out}_\lambda$ modes as

$$\varphi = \int_0^\infty d\lambda\,\left(b_\lambda\varphi^{\rm out}_\lambda + \text{h.c.}\right)\,, \tag{21}$$

and a Bogolubov transformation relates the a_ω with the b_λ. (For definitions and properties of Bogolubov transformations see Ref. [4]. We will not require these details.) To summarize, we have

$$\varphi = \int_0^\infty d\omega\,\left(a_\omega\varphi^{\rm in}_\omega + \text{h.c.}\right) = \int_0^\infty d\lambda\,\left(b_\lambda\varphi^{\rm out}_\lambda + \text{h.c.}\right)\,, \tag{22}$$

where

$$\varphi^{\rm in}_\omega \sim \frac{1}{\sqrt{4\pi\omega}}\begin{cases} e^{-i\omega v} & (\text{on } \mathcal{I}^-) \\ -e^{-i\omega G(u)} & (\text{on } \mathcal{I}^+) \end{cases}$$

$$\varphi^{\rm out}_\lambda \sim \frac{1}{\sqrt{4\pi\lambda}}\begin{cases} e^{-i\lambda u} & (\text{on } \mathcal{I}^+) \\ -e^{-i\lambda h(v)} & (\text{on } \mathcal{I}^-)\,, \end{cases} \tag{23}$$

where, if $v = G(u)$ then $u = h(v)$, and we have $G(u) = v_{\rm ent}(U(u))$ as discussed earlier.

It is now possible to write down an expression for the time dependent part of the VEV of the stress tensor at any stage during the collapse. This is given by (see Appendix A for details of the notation and derivation)

$$\langle\, T_{uu}\,\rangle^{\rm traj}_G(u) \equiv \frac{1}{12\pi}\left(\frac{dG}{du}\right)^{1/2}\partial_u^2\left(\frac{dG}{du}\right)^{-1/2} = \langle\, T_{uu}\,\rangle^{\rm ren}_G - \langle\, T_{vv}\,\rangle^{\rm ren}_G\,, \tag{24}$$

which is implicitly trajectory dependent. It can be easily verified that in the limit when the surface of the collapsing object approaches its Schwarzschild radius, the function $G(u)$ has the asymptotic form $G(u) \propto -4Me^{-u/4M}$ independent of the details of the trajectory. Hence

$$\left(\frac{dG}{du}\right)^{1/2}\partial_u^2\left(\frac{dG}{du}\right)^{-1/2} = \frac{1}{64M^2}\,, \tag{25}$$

thereby leading to a flux which asymptotes to $(\pi/12)T_H^2$ at late stages of the collapse. For later use, we write $\langle\, T_{uu}\,\rangle^{\rm traj}_G(u)$ in an alternative form. First note that since $G(u) = v_{\rm ent}(U(u))$, differentiating using the chain rule gives $G' = dG/du = (dv_{\rm ent}/dV)(dU/du) = U'/\dot{V}$ with the understanding that all functions of v such as $\dot{V} = dV/dv$ and its derivatives (see Eqn. (27) below) are to be evaluated at $v = v_{\rm ent}(U(u))$. Using the identity

$$F^{1/2}\partial_x^2 F^{-1/2} = -\frac{1}{4}\left[\,2\partial_x^2\ln F - (\partial_x \ln F)^2\,\right]\,, \tag{26}$$

that is valid for any function $F(x)$, where $\partial_x^2 \ln F = \partial_x[(\partial_x F)/F]$, and repeatedly using the chain rule, it is not hard to show that,

$$\langle T_{uu}\rangle_G^{\rm traj}(u) = \frac{1}{12\pi}\left((U')^{1/2}\partial_u^2(U')^{-1/2} - \frac{U'^2}{\dot{V}^2}(\dot{V})^{1/2}\partial_v^2(\dot{V})^{-1/2}\bigg|_{v=v_{\rm ent}(U(u))}\right)\,, \tag{27}$$

Due to the presence of the chain of functional dependences $v_{\rm ent}(U(u))$, explicit calculations of the VEV will require knowing the functions $v(V)$ and $U(u)$. In principle, once a trajectory $r = R_s(u)$ or $r = R_s(v)$ is specified, the matching conditions Eqns. (6) can be solved to get $U(u)$ and $V(v)$. In practice, it is difficult to get both $U(u)$ and $V(v)$ in closed form, and the matter is further complicated by the inversion required to get $v_{\rm ent}(V)$ at the point of entry $\mathcal{P}_{\rm ent}$. The analysis can be done for a few special cases, and Brout et al. [3, 13] for example have an explicit calculation (in parametric form) of the time-dependent flux outside a collapsing star whose internal geometry is governed by a homogeneous dust (Ref. [3], Appendix D). The asymptotic behaviour of attaining a constant flux is also displayed very nicely in that calculation (Ref. [3], Figure D.1).

B. Trajectory with null in-falling phase without horizon formation

Regardless of calculational difficulties, the general expression (24) highlights an important point : The presence of a flux of particles at large distances is governed entirely by the local dynamics (in outgoing time u) of the collapsing object. A flux will arise even in the situation where the object starts at some radius R_0 and collapses to a final smaller radius $R_f > 2M$ *without forming a black hole*, although the spectrum will not be thermal. Physically one would of course expect that the flux will only arise during the collapsing phase of the trajectory, and we will explore this feature next. We will first consider a situation in which a shell collapses from R_0 to R_f along a *null* trajectory $v =$const., after which it remains static at $r = R_f$. Although this is a rather unrealistic situation, it does admit an exact calculation of the VEV of the stress tensor at any time. We will therefore present the calculations for this case first. Following this, in the next section, we will give results for a toy trajectory which remains timelike throughout and also does not form a horizon. In this case we will be able to provide order of magnitude estimates for the second term in Eqn. (27), while the first term will be exactly calculable.

Recall that we are working in a geometry described by Eqns. (3)-(8). Our example trajectory $r = R_s$ comprises three phases –

$$\begin{aligned} u<0 &\;:\; R_s(u) = R_0 = \text{const.} \\ 0<u<u_1 &\;:\; v = \text{ const.},\; V = \text{ const.} \\ u>u_1 &\;:\; R_s(u) = 2M/(1-\epsilon^2) = \text{ const.} \end{aligned}$$

with u_1 defined in terms of ϵ as described below. We will ignore the derivative discontinuities at $u = 0$ and $u = u_1$, since these can be smoothed away if needed, and in any case will not appear when we consider the more realistic timelike trajectory in Sec. II C below. The parametrization of the final radius is chosen to ease comparison with this timelike case. For convenience we define the "tortoise" function $x(r)$ as

$$x(r) \equiv r + 2M\ln\left(r/2M - 1\right) \quad ; \quad r > 2M\,. \tag{28}$$

Using the trajectory equations (6), fixing some constants of integration and ensuring continuity at the transition events $u = 0$ and $u = u_1$, the functional form of the trajectory can be shown to be as follows: For $u < 0$ (Phase (1)), we get

$$U(u) = \left(1-\frac{2M}{R_0}\right)^{1/2} u \,-\, 2R_0 \quad ; \quad V(v) = \left(1-\frac{2M}{R_0}\right)^{1/2} v \,-\, 2x(R_0)\left(1-\frac{2M}{R_0}\right)^{1/2}\,, \tag{29}$$

For $0 < u < u_1$ (Phase (2)), we get

$$v = 2x(R_0) \quad ; \quad V = 0 \quad ; \quad U = -2R_s \quad ; \quad u = -2x(R_s) + 2x(R_0)\,, \tag{30}$$

For $u > u_1$ (Phase (3)) we get

$$U(u) = \epsilon(u-u_1) - 4M/(1-\epsilon^2) \quad ; \quad V(v) = \epsilon(v - 2x(R_0)) \quad ; \quad u_1 = 2x(R_0) - 2x(2M/(1-\epsilon^2))\,. \tag{31}$$

The Penrose diagram for this trajectory is shown in Fig. 1. There are five types of null rays (which we label I,II, ...V) relevant for the calculation of $G(u)$ and hence the stress tensor VEV, characterized by the locations of entry and exit

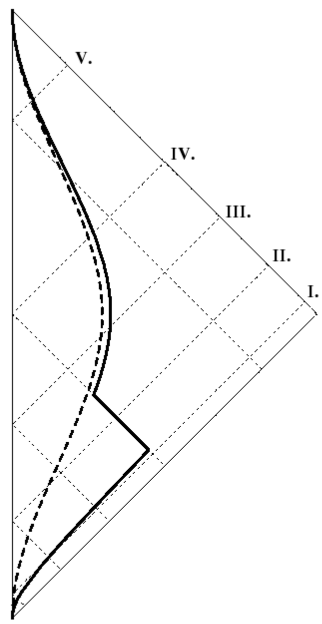


FIG. 1: Penrose diagram for the trajectory described by Eqns. (29)-(31). We have set $R_0 = 12M$ and $\epsilon^2 = 0.1$. The thick solid line is the shell trajectory, and the dashed line is the timelike surface $r = 2M$. The interior of the shell is Minkowski spacetime while the exterior is Schwarzschild. The five types of null rays relevant for the stress tensor VEV calculation are also shown.

events $\mathcal{P}_{\rm ent}$ and $\mathcal{P}_{\rm ex}$. These correspond to: (I) $\mathcal{P}_{\rm ent}$ and $\mathcal{P}_{\rm ex}$ in phase (1). (II) $\mathcal{P}_{\rm ent}$ in phase (1), $\mathcal{P}_{\rm ex}$ in phase (2). (III) $\mathcal{P}_{\rm ent}$ in phase (1), $\mathcal{P}_{\rm ex}$ in phase (3). (IV) The single ingoing ray $v = 2x(R_0)$ or $V = 0$, which skims the trajectory and finally enters the shell at $u = u_1$, with $\mathcal{P}_{\rm ex}$ in phase (3). (V) $\mathcal{P}_{\rm ent}$ and $\mathcal{P}_{\rm ex}$ in phase (3). For type (I) rays, we have

$$v_{\rm ent}(V) = \left(1 - \frac{2M}{R_0}\right)^{-1/2} V + \text{const.} \;\; ; \;\; U(u) = \left(1 - \frac{2M}{R_0}\right)^{1/2} u + \text{const.}\,, \tag{32}$$

and hence $G(u) = u+$ const. which leads to a vanishing flux. For type (II) rays, we have

$$v_{\rm ent}(V) = \left(1 - \frac{2M}{R_0}\right)^{-1/2} V + 2x(R_0) \;\; ; \;\; e^{-u/4M} = e^{(R_s - R_0)/2M}\frac{(R_s - 2M)}{(R_0 - 2M)} \;\; ; \;\; R_s = -U/2\,, \tag{33}$$

and hence

$$G(u) = \left(1 - \frac{2M}{R_0}\right)^{-1/2} U(u) + 2x(R_0)\,, \tag{34}$$

with $U(u)$ given implicitly by the last two equations in (33). We will soon return to the flux obtained from this form of $G(u)$. Note that type (II) rays all occur in $0 < u < u_1$. For type (III) rays, we have

$$v_{\rm ent}(V) = \left(1 - \frac{2M}{R_0}\right)^{-1/2} V + \text{const.} \;\; ; \;\; U(u) = \epsilon u + \text{const.}\,, \tag{35}$$

which gives $G(u) = \epsilon(1 - 2M/R_0)^{-1/2}u+$ const. which is also linear in u and hence gives a vanishing flux. Finally, type (IV) and (V) rays have

$$v_{\rm ent}(V) = \epsilon^{-1} V + \text{const.} \;\; ; \;\; U(u) = \epsilon u + \text{const.}\,, \tag{36}$$

also leading to $G(u) = u+$ const. and hence a vanishing flux.

The only nonzero flux therefore arises for type (II) rays, in the interval $0 < u < u_1$. For $\langle T_{uu} \rangle^{\rm traj}(u)$ defined in Eqn. (24), a straightforward calculation leads to

$$\langle T_{uu} \rangle^{\rm traj}(u) = \frac{1}{48\pi}\frac{M}{R_s(u)^3}\left(2 - \frac{3M}{R_s(u)}\right)\,, \quad 0 < u < u_1\,, \tag{37}$$

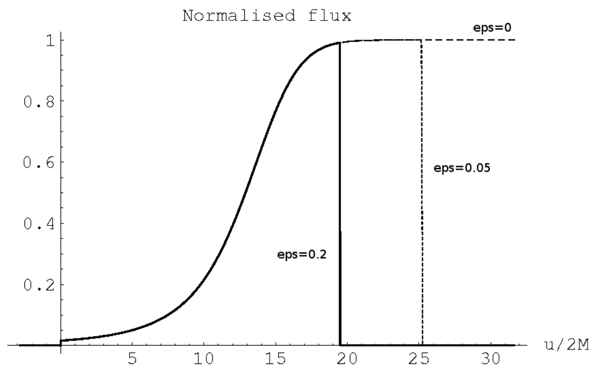


FIG. 2: The behaviour of $\langle T_{uu} \rangle^{\rm traj}(u)$ normalized by the Hawking value $(\pi/12)T_H^2$, for the trajectory studied in Sec. II B. We have set $R_0 = 12M$. The solid and dotted lines show the flux for two values of ϵ as labelled. The dashed line is the asymptotic Hawking flux corresponding to $\epsilon = 0$.

and zero otherwise, where $R_s(u)$ in the given range is implicitly determined through

$$e^{-u/4M} = e^{(R_s-R_0)/2M} \frac{(R_s - 2M)}{(R_0 - 2M)} \,. \tag{38}$$

The behaviour of $\langle T_{uu} \rangle^{\rm traj}(u)$ normalized by the Hawking value $(\pi/12)T_H^2$, is shown in Fig. 2, for $R_0 = 12M$ and two values of ϵ. The step-like rise and fall are due to the derivative discontinuities mentioned earlier, and do not occur in the timelike case which we will study next. The important feature to note is that for small *but nonzero* ϵ, the flux attains the asymptotic Hawking value but eventually falls to zero. Eqns. (37) and (38) show that the time taken to attain the Hawking value is governed by R_0, which is physically clear since the exponential redshift of the outgoing modes occurs only near $r = 2M$, so that for larger R_0, the shell spends more time radiating a smaller non-thermal flux. This argument can be easily verified numerically as well.

C. Timelike trajectory without horizon formation

We will now turn to a somewhat more realistic trajectory which is continuous, differentiable and timelike at all times, and which also does not form a horizon. The specific example we choose is a trajectory which remains fixed at $r = R_0$ until $u = 0$, and *asymptotically* (as $u \to \infty$) approaches a final radius $R_f = 2M/(1-\epsilon^2) > 2M$. It turns out that this can be easily accomplished by appropriately choosing a function $U' = dU/du$. [Since this calculation is for illustrative purposes, we will not worry about the shell dynamics needed to obtain the behaviour described below.] Consider then the following prescription for U',

$$U' = \epsilon + \left(\left(1 - \frac{2M}{R_0}\right)^{1/2} - \epsilon \right) e^{-\alpha(u)} \equiv \epsilon + A e^{-\alpha(u)} \,, \tag{39}$$

where $0 < \epsilon < (1 - 2M/R_0)^{1/2}$ is a fixed constant,

$$\alpha(u) = \frac{\theta(u)}{4M} \int_0^u d\tilde{u} h(\tilde{u}) \,, \tag{40}$$

where $\theta(u)$ is the Heaviside step function and $h(u)$ is chosen to have the following asymptotic behaviour

$$\begin{aligned} h(u) = 0 \,, u \leq 0 \;\; &; \;\; h'(u) = 0 \,, u \leq 0 \,, \\ h(u \to \infty) \to 1 \;\; &; \;\; h'(u \to \infty) \to 0 \,, \end{aligned} \tag{41}$$

and we require the asymptotic values for h and h' to be achieved exponentially fast, with a time scale determined by M. An example of a function $h(u)$ which meets these requirements is $h(u) = \theta(u) \tanh(\kappa^2 u^2)$ for some constant κ. The condition that the trajectory remain timelike throughout reduces to $-2R_s' < U'$ or $(1 - 2M/R_s) < U'$, which will in general impose a restriction on the allowed values of the timescale κ for a given starting radius R_0. Physically, if the shell starts from rest at a larger radius R_0 then it will take a longer time (at subluminal velocities throughout) to reach radii where the asymptotic exponential approach to the final radius begins. Hence a larger R_0 will imply a

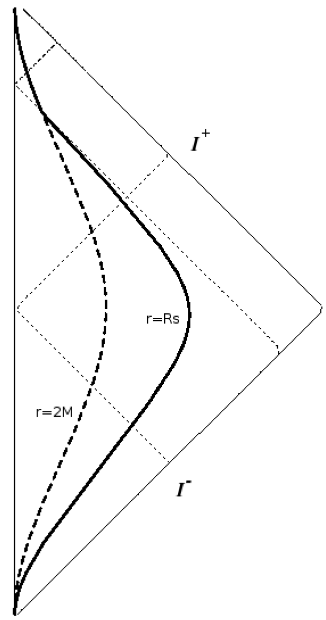


FIG. 3: Penrose diagram for the "asymptotic" trajectory described by Eqns. (39)-(41). We have set $R_0 = 5M$, $\epsilon = 10^{-2}$ and $2M\kappa = 0.075$. The thick solid line is the shell trajectory, and the dashed line is the timelike surface $r = 2M$. The interior of the shell is Minkowski spacetime while the exterior is Schwarzschild. The two null rays mark (by their entry points) the beginning of the infall phase and the approximate beginning of the asymptotic phase.

smaller κ. We have checked numerically that for $h(u) = \theta(u)\tanh(\kappa^2 u^2)$ with R_0 significantly larger than $2M$ (say $R_0 \gtrsim 4M$), the largest allowed value of κ is $\kappa_{\max} \sim M^{-1}(M/R_0)^b$ with $b \approx 2$ (and a very weak dependence on ϵ), and in general we can expect $b > 0$. The Penrose diagram for such an "asymptotic" trajectory is shown in Fig. 3.

The reason for this somewhat convoluted prescription is that we want to simplify the VEV calculations, which involve derivatives with respect to u, and it is therefore convenient to parametrize the trajectory using u. To see that the required behaviour for the trajectory $R_s(u)$ is reproduced by this U', one analyses Eqn. (7) which is reproduced below

$$2R_s'(1-U') = U'^2 - \left(1 - \frac{2M}{R_s}\right) . \tag{42}$$

The behaviour for $u \le 0$ is obtained correctly. For $u > 0$, consider the asymptotic regime (large u) where $h \approx 1$ and $h' \approx 0$. In this regime U'^2 approaches ϵ^2 from above, forcing $(1 - 2M/R_s)$ to also approach ϵ^2 asymptotically. In Appendix B 1 we show that the asymptotic behaviour for the trajectory is

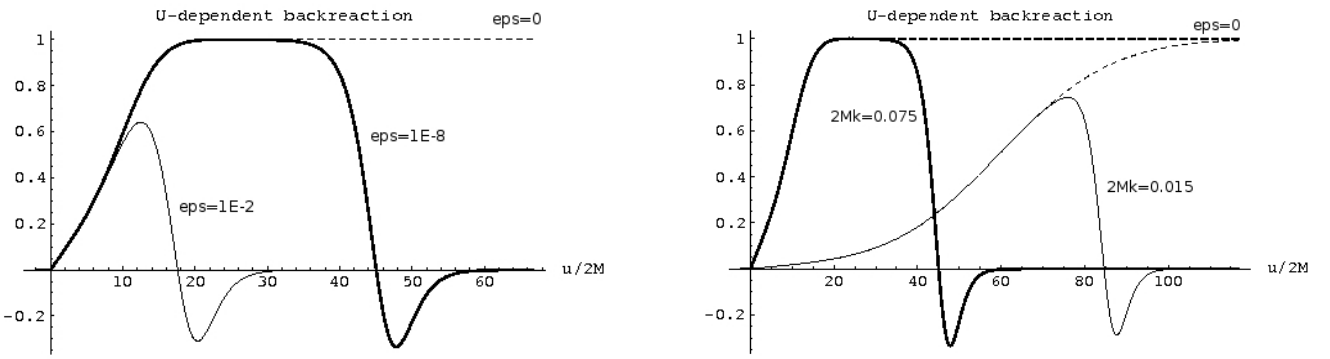
$$1 - \frac{2M}{R_s} = \epsilon^2 + f(u) \;\; ; \;\; e^{-u/4M} \ll \epsilon^2 e^{-1/\kappa M} , \tag{43}$$

where $f(u)$ is an exponentially decaying function, the exact form of which depends on the value of ϵ. Note that ϵ is fixed and we are not taking a $\epsilon \to 0$ limit. We then see that the required behaviour is being reproduced.

As before, we wish to evaluate $\langle T_{uu} \rangle_G^{\rm traj}(u)$ at any given time along this trajectory. The first term in Eqn. (27) involving derivatives of U', can be easily computed exactly using Eqn. (39) and gives

$$\frac{1}{12\pi}(U')^{1/2}\partial_u^2(U')^{-1/2} = \frac{1}{48\pi}\frac{1}{1+(\epsilon/A)e^{\alpha(u)}}\left[\alpha'^2\left(\frac{1-2(\epsilon/A)e^{\alpha(u)}}{1+(\epsilon/A)e^{\alpha(u)}}\right) + 2\alpha''\right] , \tag{44}$$

It is interesting to note that in the case $\epsilon = 0$, i.e. for a standard trajectory which approaches $r = 2M$ as $u \to \infty$, this term is positive definite and gives rise to the Hawking flux $(\pi/12)T_H^2$, which can be easily checked by setting $\epsilon = 0$ in Eqn. (44). The behaviour of $(8M)^2[(U')^{1/2}\partial_u^2(U')^{-1/2}]$ for $h(u) = \theta(u)\tanh(\kappa^2 u^2)$ with $R_0 = 5M$, is shown in Fig. 4. Panel (a) shows curves for a fixed value of κ with $2M\kappa = 0.075$ (which ensures a timelike trajectory). The two solid lines correspond to $\epsilon = 10^{-2}$ and $\epsilon = 10^{-8}$, while the dashed line corresponds to the standard case with $\epsilon = 0$, with the asymptotic value corresponding to the Hawking flux. In panel (b), the solid lines correspond to a fixed value of $\epsilon = 10^{-8}$, with $2M\kappa = 0.075$ and $2M\kappa = 0.015$. The $\epsilon = 0$ curves for these values of κ are also shown as dashed lines.



(a)Varying ϵ (b)Varying κ

FIG. 4: Behaviour of the U-dependent backreaction term $(8M)^2[(U')^{1/2}\partial_u^2(U')^{-1/2}]$, for $h(u) = \theta(u)\tanh(\kappa^2u^2)$, with $R_0 = 5M$. Panel (a) : $2M\kappa = 0.075$ is fixed. The thin solid curve has $\epsilon = 10^{-2}$, the thick solid curve has $\epsilon = 10^{-8}$, while the dashed curve is the standard case with $\epsilon = 0$. Panel (b) : For the solid curves $\epsilon = 10^{-8}$ is fixed, while the dashed curves have $\epsilon = 0$. The thick curves are for $2M\kappa = 0.075$ and the thin curves for $2M\kappa = 0.015$. See text for discussion.

We see that the initial behaviour of the U-dependent term is practically identical to the standard $\epsilon = 0$ case. Eventually the term involving α' in Eqn. (44) becomes negative, leading to negative values for the right hand side at least for the chosen $h(u)$. While this sign change might not be a generic feature, the *late time* behaviour ($u \gg \kappa^{-1}\ln(1/\epsilon^2)$, see Eqn. (43)) is generically seen to be an exponential decay $\sim (e^{1/\kappa M}/\epsilon M^2)e^{-u/4M}$, which directly follows from Eqn. (44). In Appendix B 2 we discuss some additional features of the behaviour of this U-dependent term, which allow us to place a bound on the mass loss due to backreaction for this type of trajectory.

We now turn to the second term in Eqn. (27), which involves derivatives of $\dot{V}$, and an evaluation at $v = v_{\rm ent}(U(u))$. Unfortunately, unlike Eqn. (44) for the first term, it is not possible here to give an analytic expression valid at arbitrary times. The problem arises mainly because of the function $v_{\rm ent}(V)$ which — in general — has a complicated form. Recall that this function is determined by the matching conditions on the trajectory at the point of entry $\mathcal{P}_{\rm ent}$, and the evaluation is at $V = U(u)$ where $U(u)$ is determined by the exit point $\mathcal{P}_{\rm exit}$. Despite the lack of exact expressions, it turns out that we can make order of magnitude statements by realizing that we are essentially dealing with two kinds of ingoing rays characterized by the locations of $\mathcal{P}_{\rm ent}$ and $\mathcal{P}_{\rm exit}$ on the trajectory :

1. $\mathcal{P}_{\rm ent}$ occurs for $u_{\rm ent} > 0$ but before the asymptotic phase so that $u_{\rm ent} \lesssim \kappa^{-1}\ln(1/\epsilon^2)$.
2. $\mathcal{P}_{\rm ent}$ (and hence $\mathcal{P}_{\rm exit}$) occurs in the asymptotic phase.

We ignore rays which enter at $u_{\rm ent} < 0$, since it is easy to show that $v_{\rm ent}(V) \propto V$ in this case, giving $\dot{V} =$const, and hence the second term will be exactly zero. For the choice of parameters used in Fig. 4(a) (with $\epsilon = 10^{-8}$), and fixing one constant so that $V_{\rm traj}(u = 0) = 0$, numerically one finds that the ray which enters at $u_{\rm ent} = 0$ will exit at $u_{\rm ex} \approx 14M$, and hence the second term in Eqn. (27) will vanish for $u < 14M$. In general if $V_{\rm traj}(u = 0) = 0$ then $u_{\rm ex}(u_{\rm ent} = 0)$ satisfies

$$\epsilon u_{\rm ex} + A\int_0^{u_{\rm ex}} e^{-\alpha(u)}du = 2R_0\,. \tag{45}$$

For rays of type (1), if $u_{\rm ent} \lesssim \kappa^{-1}$, the requirement that the trajectory be well-behaved in the proper frame of the shell, allows us to argue that the second term in Eqn. (27) must be small. For $\kappa^{-1} \lesssim u_{\rm ent} \lesssim \kappa^{-1}\ln(1/\epsilon^2)$, the same arguments show this term to be at most $\mathcal{O}(M^{-2})$ and changing on a timescale $\mathcal{O}(M)$ (see Appendix B 3 for details). For type (2) rays we need to work entirely in the asymptotic regime, in which the function $v_{\rm ent}(V)$ turns out to be approximately linear, and as we show in Appendix B 4 the contribution from the second term of Eqn. (27) also decays exponentially. The details depend on the value of ϵ, but the typical behaviour is

$$\frac{U'^2}{\dot{V}^2}(\dot{V})^{1/2}\partial_v^2(\dot{V})^{-1/2}\bigg|_{v=v_{\rm ent}(U(u))} \sim \frac{1}{\epsilon^2M^2}e^{1/\kappa M}e^{-\mathcal{K}u/4M} \;\;;\;\; \mathcal{K} = \mathcal{O}(1) \;\;;\;\; e^{-u/4M} \ll \epsilon^2 e^{-1/\kappa M}\,, \tag{46}$$

where $u = u_{\rm ex}$, the exit time. When ϵ is significantly smaller than unity, this term will dominate over the first term (which falls like $\sim (e^{1/\kappa M}/\epsilon M^2)e^{-u/4M}$) [22]. This analysis demonstrates a result which one should intuitively expect on physical grounds : When the object is initially static, there is no outgoing flux. When the object starts to collapse, there is a flux of at most $\mathcal{O}(M^{-2})$ which changes on a timescale $\mathcal{O}(M)$, and when the object approaches its asymptotically static configuration, this flux exponentially decays to zero.

III. THE SEMICLASSICAL BACKREACTION

The presence of an outgoing flux of particles during the gravitational collapse of a body, immediately tells us that the assumption that the mass of this body is a constant, cannot be correct. The formal way to see this is to consider the energy conservation equation $u^b T^a_{b;a} = 0$ at large distances from the collapsing object, assuming the Schwarzschild exterior (which is asymptotically flat). Taking the 4-velocity u^b to be that of an observer at rest at fixed (r, θ, ϕ), and integrating over a three volume of radius r, we have

$$\frac{dM}{dt} = \oint_r d^2S\, T^r_t = 4\pi r^2 T^r_t = {}^{(2d)}T^r_t = -\langle\, T_{uu}\,\rangle^{\rm traj}_G \equiv -L_H\,. \tag{47}$$

The first equality follows from integrating the conservation equation, with $d^2S = r^2 \sin\theta d\theta d\phi$; the second follows from spherical symmetry; the third follows from assuming that the angular components of T_{ab} vanish, so that the 4-dimensional values of T_{tt}, T_{rt} and T_{rr} are rescaled versions of their 2-dimensional counterparts; the fourth equality follows from replacing ${}^{(2d)}T_{ab}$ by its semiclassical value computed in Appendix A. The last definition is for later ease of notation.

In principle, this could lead to a situation in which the existence of an outgoing flux prevents the formation of the event horizon. The radiative flux decreases the mass of the collapsing body and thus decreases the effective radius of the event horizon which the collapsing surface is chasing. If the radius of the event horizon decreases fast enough, the collapsing surface will never be able to catch up with it and as a result the event horizon will not form. Our aim now is to analyse this situation more carefully and show that this scenario can be ruled out. Backreaction does not prevent the formation of the event horizon.

In reality, what comes to our rescue is the following fact: the mass loss rate is set by the semiclassical backreaction which is governed by the ratio $1/M^2$ in *Planckian units*, which makes it a very small number (e.g., for a solar mass object we have $M_\odot \simeq 10^{38}$). It then turns out that under the assumption that the outgoing flux (and hence mass loss rate) is small and slowly varying, a straightforward ansatz is enough to estimate the effect of the backreaction on the background geometry self-consistently. Brout et al. [3] present a calculation (which we sketch in Appendix B 5) which shows that the following late time scenario is self-consistent (see also Ref. [12], and Refs. [9, 10] for the original papers):

(a) The time dependence of the outgoing flux $L_H(u)$ at late times is determined solely by the *time-dependent* mass $M(u)$, and is given by $L_H \sim 1/M^2$, with u being an "Eddington-Finkelstein-like" outgoing coordinate. L_H is assumed to be small compared to unity (in Planckian units) and slowly varying.

(b) The exterior geometry at large distances, in terms of the outgoing coordinate u, is the outgoing Vaidya solution given by

$$ds^2_{{\rm ext,large}\,r} = -\left(1 - \frac{2M(u)}{r}\right) du^2 - 2dudr + r^2 d\Omega^2 \quad ; \quad \frac{dM(u)}{du} = -L_H\,. \tag{48}$$

(c) The exterior geometry in terms of an ingoing coordinate v, at any distance, is approximately (i.e. up to terms of order $\mathcal{O}(L_H)$) given by

$$ds^2_{\rm ext} \approx -\left(1 - \frac{2m(v,r)}{r}\right) dv^2 + 2dvdr + r^2 d\Omega^2\,, \tag{49}$$

where the mass function $m(v, r)$ is slowly varying in the entire exterior, in that

$$\frac{\partial m}{\partial v} = \mathcal{O}(L_H) \quad ; \quad \frac{\partial m}{\partial r} = \mathcal{O}(L_H)\,, \tag{50}$$

with the transformation between the u and v coordinate being such that at large distances one recovers the outgoing Vaidya metric (48) with $m(v, r) = M(u)$. (See Appendix B 5 for details of the transformation.)

(d) Self-consistency is demonstrated by showing that the VEV of the stress tensor $\langle\, T_{uu}\,\rangle^{\rm traj}(u)$ computed in this geometry does indeed behave as $\sim M(u)^{-2}(1 + \mathcal{O}(L_H))$.

The Brout et al. calculation (as acknowledged by those authors) is only valid in the regime where the flux L_H is small and slowly varying. There will inevitably be a phase in the collapse when the mass loss rate becomes significant enough that a perturbative expansion in L_H is no longer valid. A full fledged calculation of the backreaction in this regime, to our knowledge, has not yet been performed. Our interest however, is only to ask whether the backreaction

can delay the formation of the event horizon to the extent that it does not form at all. We will not worry about any significant backreaction effects that occur *after* the horizon has formed.

It is important to note that all questions about event horizon formation *must* be asked in a reference frame where this formation occurs in a finite time in the unperturbed collapse. It is not possible to theoretically settle this issue if one insists on working entirely in the coordinates used by static observers at large distances, even though these may be the most natural coordinates to use, simply because even in the *classical* scenario, event horizon formation takes an infinite amount of time in these coordinates.

Consider then a situation in which the unperturbed collapse trajectory of the shell crosses the radius $r = 2M$ (with constant M) in finite proper time with a finite subluminal velocity. The exterior geometry for the unperturbed collapse is given by the Schwarzschild metric (3), while the interior is Minkowski spacetime (5). [It is also interesting to consider a trajectory whose in-falling phase is light-like, an extension as it were of the trajectory we studied in Sec. II B. We have analysed such a trajectory in Appendix B 6.] On parametrizing the timelike trajectory using the shell proper time τ (with $ds^2|_{\rm traj} = -d\tau^2$), we find that the quantity $dv/d\tau$ remains finite at horizon formation. It is then convenient to reparametrize the trajectory using the ingoing Eddington-Finkelstein coordinate v, as $(r = \bar{R}(v), V = \bar{V}(v))$. Assuming that horizon formation in this unperturbed case occurs at some time $v = v_0$, we have

$$\bar{R}(v) - 2M = -k(v - v_0) + \mathcal{O}((v - v_0)^2) \ , \quad \text{as } v \to v_0 \, , \tag{51}$$

where the constant k can be related to the proper velocity $\beta_h \equiv -(d\bar{R}/d\tau)|_{\bar{R}=2M}$, as $k = 2\beta_h^2$ (which follows from using the metric (3) on the trajectory). The trajectory equation (8) becomes

$$2\dot{\bar{R}}(1 - \dot{\bar{V}}) = \big(1 - 2M/\bar{R}(v)\big) - \dot{\bar{V}}^2 \, , \tag{52}$$

(where the dot is a derivative with respect to v along the trajectory) which shows that $\dot{\bar{V}}$ also remains finite at $\bar{R} = 2M$.

Let us now incorporate the effects of a backreaction caused by a small outgoing flux $L_H \ll 1$, and ask by what amount is the event horizon formation delayed. Specifically, let the trajectory now be $(r = R(v), V = V(v))$ where the interior metric is still Minkowski spacetime in (r, V) coordinates (5). The exterior metric, following Brout et al., is given by Eqn. (49), and the trajectory equation (8) is easily shown to be replaced by

$$2\dot{R}(1 - \dot{V}) = (1 - 2m(v, R(v))/R(v)) - \dot{V}^2 \, . \tag{53}$$

For the exterior metric (49), the event horizon (if it forms) is the last outgoing null ray $r = r_{\rm eh}(v)$ and satisfies the outgoing null geodesic equation

$$\frac{dr_{\rm eh}}{dv} = \frac{1}{2}\left(1 - \frac{2m(v, r_{\rm eh}(v))}{r_{\rm eh}(v)}\right) \, . \tag{54}$$

Note that this event horizon is distinct from the *apparent horizon* $r = r_{\rm ah}(v)$ which is defined as the locus of events at which $dr/dv = 0$ along outgoing null geodesics, so that $r_{\rm ah}(v) = 2m(v, r_{\rm ah}(v))$. However as we show in Appendix B 5, assuming that the flux L_H is small, we get

$$2m(v, r_{\rm eh}(v)) = r_{\rm eh}(v)(1 + \mathcal{O}(L_H)) \, . \tag{55}$$

so that this distinction is irrelevant. We now use the smallness of L_H to make the following assumptions, which will turn out to be self-consistent at the end of the calculation. (a) We assume that the *functional form* of the trajectory is affected only by terms of order $\mathcal{O}(L_H)$, so that

$$R(v) = \bar{R}(v)\,(1 + \mathcal{O}(L_H)) \quad ; \quad V(v) = \bar{V}(v)\,(1 + \mathcal{O}(L_H)) \ , \tag{56}$$

which is reasonable since the trajectory of the event horizon ($r = 2M$ in the unperturbed case) is also affected by the same amount (see Eqns. (54), (55)); and (b) the event horizon is formed (i.e. $R(v) = r_{\rm eh}(v)$ is satisfied) at some finite $v = \tilde{v}_0$ such that $\tilde{v}_0 = v_0 + \delta v$ where $\bar{R}(v_0) = 2M$ and δv is to be determined. We now evaluate the perturbed trajectory equation (53) at $v = \tilde{v}_0$, linearize around $v = v_0$ assuming that δv is "small", and use the unperturbed equation (52) to obtain after straightforward algebra,

$$\delta v = \mathcal{O}\left(\frac{2M}{k} L_H\right) = \mathcal{O}\left(\frac{M}{\beta_h^2} L_H\right) \, , \tag{57}$$

where β_h is the proper velocity with which the unperturbed trajectory crosses $r = 2M$ and L_H is the luminosity of the thermal radiation. This result is intuitively clear : The delay in the event horizon formation is governed by

the product of L_H and the timescale of the unperturbed collapse, which is essentially set by the initial mass of the collapsing object.

One might still argue however, that there could in principle exist *unperturbed* trajectories which slow down sufficiently while still outside $r = 2M$, that β_h becomes significantly smaller than unity. In such a case, it might be possible to delay the event horizon formation to such an extent that the backreaction itself starts changing significantly and becomes large, leading to a runaway process which might exclude the formation of the event horizon. While this situation cannot be excluded, the following argument does render it implausible.

- First note that there is no semiclassical radiation in static geometries. The flux of particles arises when the geometry (characterized e.g. by the shell radius) changes by a significant amount during the time that an ingoing mode reflects at the center and exits the object. For this to occur the trajectory must be governed by a timescale *linear* in M, since this is the order of the time spent by the modes inside the object. For example in the "asymptotic" trajectory of Sec. II C, the maximum flux is achieved for times $u \gg \kappa^{-1}$ when the timescale of the function $\alpha(u)$ is set by M.

- Our calculations in Sec. II, and also e.g. the calculation in Brout et al.'s Appendix D, show that in any stage of the trajectory governed by the timescale M, the flux is expected to be at most $\mathcal{O}(1/M^2)$, with the largest value expected only as the object approaches $r = 2M$. In Sec. II C we also saw that in the extreme case when the trajectory is asymptotically slowed down to a halt, the flux in fact exponentially decays. One expects therefore that if a trajectory is slowed down by some smaller amount (i.e. not exponentially), the flux will have to lie between its maximum of $\mathcal{O}(1/M^2)$, and zero. In particular, slowing down a trajectory cannot *increase* the flux.

- In order to completely evaporate the collapsing object before the event horizon forms, what we need along the *unperturbed* trajectory is to sustain the maximum possible flux for the largest possible time. Naively one would want a flux $\mathcal{O}(1/M^2)$ (which is the maximum possible) sustained for a time $\mathcal{O}(M^3)$. The previous arguments show that this is unlikely to happen, since there is a trade-off between having a significant flux and remaining outside $r = 2M$ for long enough.

A. Bound on the total radiated mass

For trajectories such as those discussed in Sec. II, and in fact for any trajectories that have such a form up to any finite time, we can place concrete bounds on the amount of mass that can be radiated away. We do this by calculating or estimating the integral $\Delta M = \int \langle T_{uu} \rangle^{\rm traj} du$. For the trajectory of Sec. II B, $\langle T_{uu} \rangle^{\rm traj}$ is non-zero only in the range $0 < u < u_1$, in which its integral can be explicitly performed to give

$$\frac{\Delta M}{M} = \frac{-1}{48\pi}\frac{1}{4M^2}\left[\ln(\epsilon^{-2}) + \ln(1 - 2M/R_0) + (2M/R_0) - (1 - \epsilon^2) + \frac{3}{2}\left\{(2M/R_0)^2 - (1 - \epsilon^2)^2\right\}\right] . \tag{58}$$

We expect R_0/M to be significantly larger than unity. Also, since we are in the semiclassical domain, we can only expect to make meaningful statements for situations where the final asymptotic radius is *not closer to* $r = 2M$ than one Planck length. This Planck length bound translates to a lower bound on ϵ given by $\epsilon^2 > 1/(2M + 1)$, so that the maximum mass that can be radiated away (ignoring numerical factors) is

$$|\Delta M_{\rm max}| \sim M^{-1} \ln(M) . \tag{59}$$

For the asymptotic trajectory of Sec. II C, the situation is trickier since we do not have exact expressions for the full backreaction, and also because the presence of the timescale κ^{-1} causes a subtle interplay of competing effects. We show in Appendix B 2 however, that even in this case there is an upper bound on the possible mass loss, which is in fact *identical* (up to numerical factors) to the expression (59).

IV. CONCLUSIONS

Since we have provided a detailed summary of results right at the beginning of the paper, we shall be brief in this section and will just review and stress the key conclusions which we have obtained.

It is possible to explicitly compute the VEV of the stress tensor around a collapsing shell for any arbitrary trajectory of the shell. In principle, this calculation should be done in (1+3) dimensions but the results based on s-waves can be obtained by using conformal field theory techniques in the dimensionally reduced (1+1) case. Such an analysis shows that the time dependent geometry around a collapsing shell will lead to emission of a flux of particles to infinity though the spectrum, in general, will not be thermal. This result has nothing to do with black holes or event horizons.

When the shell does collapse to a radius close to the Schwarzschild radius, the relevant functions acquire the well known universal form characteristic of an exponential redshift of the outgoing modes. This exponential redshift leads to a thermal spectrum. In other words while any collapsing shell will lead to a radiation flux, such a flux will not, in general, have any universal characteristic. But if the collapsing shell goes close to its Schwarzschild radius, the flux becomes approximately thermal.

If the collapsing shell follows a trajectory which does not form an event horizon and — instead — asymptotes to a radius $2M(1-\epsilon^2)^{-1}$, then we still get (approximately) thermal radiation during the period $M \lesssim t \lesssim M\ln(1/\epsilon^2)$. Of course, in the case $\epsilon = 0$, the collapse leads to an event horizon and we get the standard result that thermal radiation will be emitted at all late times corresponding to $t \gtrsim M$. On the other hand, for $\epsilon \neq 0$, at $t > M\ln(1/\epsilon^2)$ the radiation decays to zero exponentially as to be expected around a final static configuration.

The above description does not take into account the effect of backreaction on the formation of event horizon. If radiation of energy can decrease the mass (and hence the effective radius of the event horizon) at a sufficiently rapid pace, then the collapsing shell might never catch up with the event horizon and a black hole may never form. This, however, does not happen essentially because the amount of radiation emitted during the collapse is not enough to decrease the effective radius of the event horizon sufficiently fast. This result is nontrivial and requires the careful computations which we have performed in this paper.

Acknowledgments

We thank T. P. Singh and Dawood Kothawala for helpful discussions. TP gratefully acknowledges hospitality at TIFR during a visit in which this work was begun. AP gratefully acknowledges hospitality at IUCAA during a visit in which a part of this work was completed.

APPENDIX A: VEV OF STRESS TENSOR AND RELATED ISSUES

In this appendix we present a pedagogical review of the calculation of the VEV of the stress tensor, using the tools of 1+1 conformal field theory, and highlight certain conceptual issues. The reader is referred to Refs. [3, 4, 8, 11] for details of the original calculations. The notation used here was laid out in Sec. II A.

1. Choice of vacuum state

In the Heisenberg picture, there is only one vacuum state that one can work with, which is the $|\texttt{in}\rangle$ state. Nevertheless, it is common to find references in the literature to vacua defined with respect to various combinations of ingoing and outgoing null coordinates (for a review of these vacuum states see Ref. [23]). One particular vacuum known as the Unruh vacuum takes on a special significance in the collapse situation. This is the vacuum defined to be positive frequency with respect to the ingoing Eddington-Finkelstein coordinate v and the outgoing *Kruskal* coordinate $\mathcal{U} = -4Me^{-u/4M}$ [11]. As the surface of the object approaches its Schwarzschild radius, one finds that under the generic assumption that the trajectory in the (v,r) coordinates remains well-behaved at the Schwarzschild radius, we get $G(u) \to -4Me^{-u/4M} = \mathcal{U}$ as $u \to \infty$. This behaviour arises due to the presence of an ingoing ray labelled by $v = v_H$ say, which after reflection is the *last ray* that can exit the object before the horizon forms (i.e. before the trajectory crosses $r = 2M$). The presence of this "last escaper" ingoing ray means that in the large u limit for the outgoing rays, the corresponding ingoing rays bunch together around $v \lesssim v_H$, and hence the function $v_{\rm ent}(V)$ can be linearized. Additionally, assuming that the trajectory crosses the Schwarzschild radius at finite $v = v_0$ with a finite non-zero velocity $dR_s/dv|_{v_0}$, it is straightforward to show using Eqns. (6) that as $u \to \infty$ (i.e. $R_s \to 2M$), $U(u) \sim e^{-u/4M} + \mathcal{O}(e^{-u/2M})$. This shows that at late times $G(u) \propto e^{-u/4M}$, and in this limit, the G-vacuum is identical to the Unruh vacuum.

The previous arguments provide a dynamical "reason" for "choosing" the Unruh vacuum at late times : the reason is simply the Heisenberg picture statement that the state does not evolve. The rest is taken care of by the dynamics and boundary conditions.

2. Stress tensor VEVs : Regularization

The object of interest is $\langle\texttt{in}|T_{ab}|\texttt{in}\rangle$, where T_{ab} is the *2-dimensional* stress tensor in the $u-v$ plane. Provided the angular components of the stress tensor vanish, the four dimensional stress tensor can be obtained simply by rescaling the 2-d one by $(4\pi r^2)^{-1}$. The classical value of T_{ab} for a 2-d massless scalar field is $\partial_a\varphi\partial_b\varphi - (1/2)g_{ab}\partial^i\varphi\partial_i\varphi$, and

it is then easy to see by expanding the field operator in the modes $\varphi_\lambda^{\rm out}$, that $\langle \texttt{in}|T_{ab}|\texttt{in}\rangle$ corresponds ultimately to expectation values of the number operators $b_\lambda^\dagger b_\lambda$ in the G-vacuum. Operationally one works with an expansion in the $\varphi_\omega^{\rm in}$ modes since the state $|\texttt{in}\rangle$ is the vacuum of the G-modes.

However, the stress tensor VEV needs to be regularized, since even in Minkowski spacetime with $\sqrt{4\pi\omega}\varphi_\omega = e^{-i\omega v} - e^{-i\omega u}$, this object diverges. So what we want is

$$\langle T_{ab} \rangle^{\rm ren} = \langle \texttt{in}|T_{ab}|\texttt{in}\rangle - \text{“Minkowski value”} \,. \tag{A1}$$

a. The accelerating mirror

Let us first consider the toy example of a scalar field quantized in flat spacetime ($ds^2 = -dudv + ...$) with the boundary condition provided by an accelerating mirror which follows the trajectory $v = G(u)$ for some $G(u)$. Note that since there is no “interior” region in this problem, the complication of finding a function like $v_{\rm ent}(V)$ does not exist. The analysis is therefore simpler, and will help to set up some formalism which can then be easily adapted to the case of a genuine collapse situation. In particular one finds that the renormalized VEV required can ultimately be obtained in a fairly straightforward manner, without going into details of contour integrals etc. While the interesting case occurs when $G(u) = -(1/a)e^{-au}$ for constant a, in which case reflection from the mirror leads to a thermal spectrum of particles in the G-vacuum, we will not require this form of G in the calculation below. The flat spacetime mode functions satisfying the reflection condition on $v = G(u)$ are,

$$f_\omega = \frac{1}{\sqrt{4\pi\omega}}\left(e^{-i\omega v} - e^{-i\omega G(u)}\right) . \tag{A2}$$

The “Minkowski value” for the stress tensor component T_{vv} is simply $\langle \texttt{in}|T_{vv}|\texttt{in}\rangle$, while for T_{uu} it is $\langle \texttt{out}|T_{uu}|\texttt{out}\rangle$. This follows from the fact that the Minkowski vacuum has modes which are positive frequency with respect to u and v. This gives us

$$\langle T_{vv} \rangle^{\rm ren} = 0 \;\; ; \;\; \langle T_{uu} \rangle^{\rm ren} = \frac{1}{12\pi}(G')^{1/2}\partial_u^2 (G')^{-1/2} \,, \tag{A3}$$

where $G' \equiv dG/du$. When $G' = e^{-au}$, the second expression can be obtained by explicitly writing out the mode expansions and performing the integrals involved (see Ref. [3], Sec. 2.5.2). In general, the following regularisation procedure proves to be very useful, since it carries over to the gravitational case as well (see Ref. [4] for alternative regularisation procedures).

One notes that the u-part of the Green function for the field in the u-vacuum $|\texttt{out}\rangle$, is $(-1/4\pi)\ln|u-u'|$ where one is considering two spacetime events (u,v) and (u',v'). Similarly, the Green function in the G-vacuum $|\texttt{in}\rangle$ is $(-1/4\pi)\ln|G(u) - G(u')|$. The stress tensor VEV in the state $|\texttt{in}\rangle$ can then be computed using a point split double derivative, as

$$\begin{aligned}\langle \texttt{in}|T_{uu}|\texttt{in}\rangle &= \lim_{u\to u'} \langle \texttt{in}|\partial_u\varphi(u)\partial_{u'}\varphi(u')|\texttt{in}\rangle \\ &= \lim_{u\to u'} \partial_u\partial_{u'}\langle \texttt{in}|\varphi(u)\varphi(u')|\texttt{in}\rangle \\ &= \lim_{u\to u'} \partial_u\partial_{u'}\left(-\frac{1}{4\pi}\ln|G(u) - G(u')|\right) \,,\end{aligned} \tag{A4}$$

and similarly $\langle \texttt{out}|T_{uu}|\texttt{out}\rangle$, which gives the result in Eqn. (A3) after expanding the expressions to third order in $(u-u')$. In order to get the result in a form which is generalizable to curved spacetime as well, one notes that the above argument for calculating the uu components can be inverted to get the GG components. One would do this by taking a point split double derivative with respect to G instead of u, and get

$$\begin{aligned}\langle T_{GG} \rangle^{\rm ren} = \langle \texttt{in}|T_{GG}|\texttt{in}\rangle - \langle \texttt{out}|T_{GG}|\texttt{out}\rangle &= -\frac{1}{4\pi}\lim_{G\to G'}\partial_G\partial_{G'}\left(\ln|G-G'| - \ln|u(G) - u(G')|\right) \\ &= -\frac{1}{12\pi}\left(\frac{du}{dG}\right)^{1/2}\partial_G^2\left(\frac{du}{dG}\right)^{-1/2} \,.\end{aligned} \tag{A5}$$

It is important to note that the u coordinate here corresponds to the *Minkowski* u.

b. Genuine gravitational fields

Let us now turn to the original problem of a body collapsing to form a black hole. Due to mathematical similarities with the moving mirror problem, one can use similar ideas here as well, with some subtle differences [3, 8]. We still want a regularized version of $\langle \mathtt{in}|T_{ab}|\mathtt{in}\rangle$, the 2-dimensional stress tensor VEV. But now, there is no globally defined "Minkowski value" which one can subtract. Instead, one subtracts a *locally defined* object corresponding to the stress tensor VEV in a *locally inertial vacuum*. Given any spacetime with

$$ds^2 = -C(G,v)dGdv + \dots , \tag{A6}$$

such that we want expectation values with respect to the vacuum of the modes $e^{-i\omega v}$ and $e^{-i\omega G}$, one can show that locally inertial coordinates are given, at some event $x \equiv (G_0, v_0)$, by

$$\hat{u} = \int_{G_0}^{G} \frac{C(G', v_0)}{C(G_0, v_0)} dG' \;\; ; \;\; \hat{v} = \int_{v_0}^{v} \frac{C(G_0, v')}{C(G_0, v_0)} dv' \, . \tag{A7}$$

$\hat{u}$ and $\hat{v}$ are affine parameters along the radial null geodesics $v = v_0$ and $G = G_0$ respectively, which pass through x, and the Christoffels in the $(\hat{u}, \hat{v})$ coordinates vanish at x [3]. At this event x one wants to subtract from $\langle \mathtt{in}|T_{ab}|\mathtt{in}\rangle$, the quantity $\langle \mathtt{I(x)}|T_{ab}|\mathtt{I(x)}\rangle$ where $|\mathtt{I(x)}\rangle$ is the vacuum corresponding to the "locally inertial modes" $e^{-i\omega\hat{u}}$, $e^{-i\omega\hat{v}}$ at x.

The mathematics is now identical to the moving mirror case, and one finds

$$\begin{aligned} \langle T_{GG} \rangle_G^{\mathrm{ren}} &= -\frac{1}{12\pi}\left(\frac{d\hat{u}}{dG}\right)^{1/2} \partial_G^2 \left(\frac{d\hat{u}}{dG}\right)^{-1/2} , \\ \langle T_{vv} \rangle_G^{\mathrm{ren}} &= -\frac{1}{12\pi}\left(\frac{d\hat{v}}{dv}\right)^{1/2} \partial_G^2 \left(\frac{d\hat{v}}{dv}\right)^{-1/2} , \end{aligned} \tag{A8}$$

where the additional subscript on the VEVs reminds us that we are working in the G-vacuum (which is of course also the v-vacuum). Eqn. (A7) can be used to remove all explicit reference to the locally inertial coordinates, and we find in general that

$$\langle T_{GG} \rangle_G^{\mathrm{ren}} = -\frac{1}{12\pi} C^{1/2} \partial_G^2 C^{-1/2} \;\; ; \;\; \langle T_{vv} \rangle_G^{\mathrm{ren}} = -\frac{1}{12\pi} C^{1/2} \partial_v^2 C^{-1/2} \, . \tag{A9}$$

For our case, the conformal factor is given in the exterior Schwarzschild geometry, by

$$C = \left(1 - \frac{2M}{r}\right) \frac{du}{dG} \, , \tag{A10}$$

which follows from comparing the metrics in Eqns. (3) and (A6), where u is the outgoing Eddington-Finkelstein coordinate. This simple prescription for constructing the stress tensor VEV is a consequence of considering only the s-wave contribution and further dropping the Schwarzschild potential barrier, which is what allowed us to approximate the scalar field modes in the exterior as propagating on a conformally flat geometry. See Ref. [3] for a discussion on how good this approximation is in practice.

3. Changing vacua

Let us see how the renormalized stress tensor VEV behaves under a change of the reference vacuum. This will allow us to write expressions which will clarify what happens in various physical situations. The structure of this object is, e.g.

$$\langle T_{GG} \rangle_G^{\mathrm{ren}} \sim C^{1/2} \partial_G^2 C^{-1/2} \sim \underbrace{\partial_G \partial_{G'}}_{\text{Choice of coords}} \left[\ln \underbrace{|G - G'|}_{\text{Choice of vacuum}} - \ln |\hat{u} - \hat{u}'| \right] \tag{A11}$$

This shows that in any given vacuum, changing coordinates is trivial, one simply multiplies by the appropriate coordinate transformation factors. On the other hand, in a given set of coordinates, changing the *vacuum* is not as trivial, although it is straightforward. Let us suppose that we want to compute $\langle T_{ab} \rangle$ in the vacuum corresponding to some null coordinates (f, g) instead of (G, v), where we assume that $f = f(G)$, $g = g(v)$ so that the 2-d metric

remains conformally flat. [Later we will specialize to the case $f = u$, $g = v$.] Our previous results (Eqn. (A9)) tell us that

$$\langle T_{ff} \rangle_{fg}^{ren} = -\frac{1}{12\pi}(C_{fg})^{1/2}\partial_f^2(C_{fg})^{-1/2} \quad ; \quad \langle T_{gg} \rangle_{fg}^{ren} = -\frac{1}{12\pi}(C_{fg})^{1/2}\partial_g^2(C_{fg})^{-1/2} \,, \tag{A12}$$

with the subscript $\langle \, \rangle_{fg}$ on the VEVs denoting the choice of vacuum, and where

$$C_{fg} = C\left(G(f), v(g)\right)\frac{dG}{df}(f)\frac{dv}{dg}(g) \,. \tag{A13}$$

In general for some function $F(x)$, we have the identity

$$F^{1/2}\partial_x^2 F^{-1/2} = -\frac{1}{4}\left[\, 2\partial_x^2 \ln F - (\partial_x \ln F)^2 \,\right] \,. \tag{A14}$$

Using this and the expression for C_{fg} in Eqn. (A13), one can simplify Eqn. (A12) to obtain

$$\begin{aligned} \langle T_{ff} \rangle_{fg}^{\rm ren} &= \left(\frac{dG}{df}\right)^2 \langle T_{GG} \rangle_{G}^{\rm ren} - \frac{1}{12\pi}\left(\frac{dG}{df}\right)^{1/2} \partial_f^2 \left(\frac{dG}{df}\right)^{-1/2} \,, \\ \langle T_{gg} \rangle_{fg}^{\rm ren} &= \left(\frac{dv}{dg}\right)^2 \langle T_{vv} \rangle_{G}^{\rm ren} - \frac{1}{12\pi}\left(\frac{dv}{dg}\right)^{1/2} \partial_g^2 \left(\frac{dv}{dg}\right)^{-1/2} \end{aligned} \tag{A15}$$

In particular, for the "Boulware" vacuum defined through the modes $e^{-i\omega u}$ and $e^{-i\omega v}$, we have $f = u$, $g = v$. The Boulware vacuum attains physical relevance only in the static case when the star/shell does not evolve, and one has $G(u) = u+$const (which can be shown using the trajectory and matching equations using $R_s =$const). But the vacuum itself can of course always be defined even in the dynamical situation. For this vacuum,

$$\begin{aligned} \langle T_{uu} \rangle_{B}^{\rm ren} &= \langle T_{uu} \rangle_{G}^{\rm ren} - \frac{1}{12\pi}\left(\frac{dG}{du}\right)^{1/2} \partial_u^2 \left(\frac{dG}{du}\right)^{-1/2} \,, \\ \langle T_{vv} \rangle_{B}^{\rm ren} &= \langle T_{vv} \rangle_{G}^{\rm ren} \,. \end{aligned} \tag{A16}$$

As discussed earlier, in the limit when the object approaches $r = 2M$, we have $G(u) \to -4Me^{-u/4M}$, and it is not hard to show that

$$\langle T_{uu} \rangle_{B}^{\rm ren} = \langle T_{uu} \rangle_{\mathcal{U}}^{\rm ren} - \frac{\pi}{12}T_H^2 \,, \tag{A17}$$

where $T_H = 1/(8\pi M)$ is the Hawking temperature of the collapsing object, and we have denoted the asymptotic G-vacuum by the subscript $\mathcal{U}$ (for $\mathcal{U}$nruh). Using the fact that the conformal factor for the Boulware vacuum is simply $(1 - 2M/r)$, where r is given implicitly in terms of u and v by the relation (4), the stress tensor VEV in the Boulware vacuum can be explicitly computed as

$$\langle T_{uu} \rangle_{B}^{\rm ren} = -\frac{1}{48\pi}\frac{M}{r^3}\left(2 - \frac{3M}{r}\right) \,. \tag{A18}$$

This gives us the relations

$$\langle T_{uu} \rangle_{B}^{\rm ren}(r) = -\frac{\pi}{12}T_H^2\left[\frac{32M^3}{r^3} - \frac{48M^4}{r^4}\right] = \langle T_{vv} \rangle_{B}^{\rm ren} = \langle T_{vv} \rangle_{\mathcal{U}}^{\rm ren} \,, \tag{A19}$$

$$\begin{aligned} \langle T_{uu} \rangle_{\mathcal{U}}^{\rm ren}(r) &= \langle T_{uu} \rangle_{B}^{\rm ren}(r) - \langle T_{uu} \rangle_{B}^{\rm ren}(r = 2M) \\ &= \frac{\pi}{12}T_H^2\left[1 - \frac{32M^3}{r^3} + \frac{48M^4}{r^4}\right] \\ &= \frac{\pi}{12}T_H^2\left(1 - \frac{2M}{r}\right)^2\left[1 + \frac{4M}{r} + \frac{12M^2}{r^2}\right] \,, \end{aligned} \tag{A20}$$

where $\langle T_{uu} \rangle_{B}^{\rm ren} = \langle T_{vv} \rangle_{B}^{\rm ren}$ follows from the fact that $\partial_u r = -\partial_v r$. We see that $\langle T_{uu} \rangle_{\mathcal{U}}^{\rm ren}$ vanishes on $r = 2M$ but attains the constant Hawking flux value as $r \to \infty$. Also, since the coordinate transformation factor relevant to the u

	$G \to \mathcal{U} \to 0$; $r = $const.$< \infty$ Static asymptotic observer at finite $r > 2M$	$G \to \mathcal{U} \to 0$; $r \to \infty$ Static asymptotic observer at spatial infinity	$G \to \mathcal{U} \to 0$; $(r - 2M) \propto \mathcal{U}$ Freely falling asymptotic observer
$\langle T_{uu} \rangle_B^{\text{ren}}$	Eqn. (A19)	0	$-(\pi/12)T_H^2$
$\langle T_{\mathcal{U}\mathcal{U}} \rangle_B^{\text{ren}}$	∞	0 (if $r \to \infty$ first)	$-\infty$
$\langle T_{uu} \rangle_{\mathcal{U}}^{\text{ren}}$	* Eqn. (A20)	* $(\pi/12)T_H^2$	0
$\langle T_{\mathcal{U}\mathcal{U}} \rangle_{\mathcal{U}}^{\text{ren}}$	∞	∞	* Finite, obs. dep.

TABLE I: Various limiting cases of the stress tensor VEV in different vacua.

sector is $d\mathcal{U}/du = (-\mathcal{U}/4M)$ which goes to zero in the asymptotic limit, we see that the component $\langle T_{\mathcal{U}\mathcal{U}} \rangle_B^{\text{ren}}(r) \to \infty$ in this asymptotic limit, for *any finite value* of r. [The limit $\mathcal{U} \to 0$ is distinct from the limits $r \to 2M$ or $r \to \infty$.]

The object $\langle T_{\mathcal{U}\mathcal{U}} \rangle_{\mathcal{U}}^{\text{ren}}$ is trickier. This object has a factor $(4M/\mathcal{U})^2$ *and* a factor $(1 - 2M/r)^2$ in the asymptotic limit. Clearly it remains finite *on the horizon* if one takes the $\mathcal{U} \to 0$ and $r \to 2M$ limits simultaneously, assuming $\mathcal{U} \propto (r - 2M)$. Such a situation is relevant for an *asymptotic freely falling observer* maintaining approximately constant v, for example a geodesic observer who starts from rest at some very large distance. This result therefore shows that such an observer must see a finite flux at late times, although apparently this flux is not universal and will depend on the specific trajectory of the freely falling observer.

For $r = $ const. $> 2M$, the Kruskal coordinates are not the natural choice; such an observer uses the Eddington-Finkelstein (u, v) null coordinates. Hence, even though $\langle T_{\mathcal{U}\mathcal{U}} \rangle_{\mathcal{U}}^{\text{ren}}$ diverges as $\mathcal{U} \to 0$ for fixed r, this is not physically relevant. The natural choice for such an observer in the $\mathcal{U} \to 0$ limit is $\langle T_{uu} \rangle_{\mathcal{U}}^{\text{ren}}(r)$, which as we see remains finite and attains the Hawking flux value at large distances $r \to \infty$. Table I summarizes these results. The boxes highlighted by asterisks show physically relevant limits. Not surprisingly, all of these are finite.

4. Stress tensor VEV at arbitrary times

The calculations in the previous section show that in the limit when the collapsing object approaches its Schwarzschild radius, the stress tensor VEV becomes time independent and takes the form given in Eqns. (A20) and (A19) for the components $\langle T_{uu} \rangle_G^{\text{ren}}$ and $\langle T_{vv} \rangle_G^{\text{ren}}$ respectively, in the limit $G \to \mathcal{U}$. Physically this corresponds to a steady blackbody flux of particles at any radius r given by $T_t^r = T_{vv} - T_{uu} = -(\pi/12)T_H^2$. Showing that this flux is associated with a thermal spectrum requires a calculation of the Bogolubov coefficients relating the Boulware and Unruh vacua (see e.g. Ref. [3]). We will not discuss this calculation since we wish to discuss backreaction, for which we need the flux and not the explicit form of the particle spectrum. In principle though, one already has the form of the stress tensor VEV at *any* stage during the collapse, since we have derived the relation (see Eqns. (A16) and (A19))

$$\langle T_{uu} \rangle_G^{\text{ren}} - \langle T_{vv} \rangle_G^{\text{ren}}(r) = \frac{1}{12\pi} \left(\frac{dG}{du}\right)^{1/2} \partial_u^2 \left(\frac{dG}{du}\right)^{-1/2} \equiv \langle T_{uu} \rangle_G^{\text{traj}}(u) \,, \tag{A21}$$

which defines the trajectory-dependent quantity $\langle T_{uu} \rangle_G^{\text{traj}}(u)$ which asymptotes to $(\pi/12)T_H^2$ at late stages of the collapse. Here $\langle T_{vv} \rangle_G^{\text{ren}}(r) = \langle T_{uu} \rangle_B^{\text{ren}}(r)$ is given by Eqn. (A18) or Eqn. (A19).

APPENDIX B: PROOFS FOR VARIOUS RESULTS

In this appendix we give calculational details for some of the results quoted in the main text.

1. Asymptotic behaviour of the trajectory defined by Eqn. (39)

As mentioned in the text, for U' given by Eqn. (39), the trajectory equation (42) forces $(1 - 2M/R_s)$ to approach ϵ^2 at late times. Let us assume therefore, that for large enough times the asymptotic behaviour is given by

$$1 - \frac{2M}{R_s} = \epsilon^2 + f(u) \quad ; \quad |f(u)| \ll \epsilon^2 \,. \tag{B1}$$

We will soon determine the condition to be satisfied by u in order for this to hold. Differentiating and rewriting Eqn. (B1) leads to

$$2R_s' = \left(\frac{R_s}{2M}\right)^2 4Mf' \quad ; \quad \frac{R_s}{2M} = \frac{1}{1-\epsilon^2-f}\,, \tag{B2}$$

which are exact. The behaviour of U' in the asymptotic regime is well approximated by $U' = \epsilon^2 + \tilde{A}e^{-u/4M}$, where $\tilde{A}$ is different from the A which appears in Eqn. (39). The assumption that the asymptotic behaviour for the function $h(u)$ in Eqn. (41) is achieved exponentially on a timescale κ^{-1}, means that we can expect $\tilde{A} = \mathcal{O}(e^{1/\kappa M})$. Recall from the main text that we have

$$\kappa < \kappa_{\max} \sim M^{-1}(M/R_0)^b \quad ; \quad b > 0\,. \tag{B3}$$

Using the expressions in Eqn. (B2) in the trajectory equation (42) and linearising in f leads to

$$4Mf'(1-\epsilon) \approx (1-\epsilon^2)^2\left(2\epsilon\tilde{A}e^{-u/4M} - f\right)\,. \tag{B4}$$

Defining the variable $s \equiv u/4M$, we have

$$\partial_s f + K_1 f = K_2 e^{-s}\,, \tag{B5}$$

where

$$K_1 \equiv \frac{(1-\epsilon^2)^2}{(1-\epsilon)} \quad ; \quad K_2 = 2\epsilon\tilde{A}K_1\,. \tag{B6}$$

Solving Eqn. (B5) gives

$$f(u) = Ce^{-K_1 u/4M} + \frac{K_2}{K_1-1}e^{-u/4M} \quad ; \quad \text{if } K_1 \neq 1\,, \tag{B7}$$

with $f = (C + K_2 u/4M)e^{-u/4M}$ when $K_1 = 1$, which we ignore since this solution forms a set of measure zero. The condition $|f| \ll \epsilon^2$ shows that for consistency we must have

$$u \gg \kappa^{-1}\ln(1/\epsilon^2)\,, \tag{B8}$$

which follows from the fact that for small ϵ, $K_2/(K_1-1) \sim \tilde{A} \sim e^{1/\kappa M}$. The constant C is not constrained by this analysis : it can only be fixed by evolving the full set of equations. The best we can say is that it must be at most of order $\sim K_2/(K_1-1)$, to be consistent with the condition $|f| \ll \epsilon^2$. Consequently the sign of f is also not fixed by this analysis. We see that $f(u)$ exponentially decays and hence consistently reproduces $(1-2M/R_s) \to \epsilon^2$ at late times.

2. Behaviour of the U-dependent backreaction and bound on mass loss

In this section we discuss some features of the backreaction term $(U')^{1/2}\partial_u^2(U')^{-1/2}$, and derive an upper bound on the expected mass loss due to backreaction in the "asymptotic" trajectory of Sec. II C.

a. Analysing Fig. 4

An important feature in Fig. 4 is the time ($u = u_*$ say) at which any given curve starts deviating from the $\epsilon = 0$ case. This time depends not only on ϵ but also on κ. (For now we ignore the dependence on the starting radius R_0 and assume $\kappa < \kappa_{\max}$.) We see that for a given κ, smaller values of ϵ will lead to a larger u_*/M, which can be understood as the fact that it will take longer for a trajectory to "feel the effect" of a smaller ϵ.

On the other hand, for a given ϵ, changing the value of κ has three distinct effects.

- Firstly, reducing κ "stretches out" the *standard* $\epsilon = 0$ behaviour of the backreaction, which is expected since the timescale over which the standard trajectory approaches $r = 2M$ is governed by κ^{-1} which is now larger.
- Secondly, reducing κ increases the value of u_*/M at which the backreaction deviates from the standard case, which is also expected since a larger κ^{-1} will slow down the rate at which $\alpha(u)$ increases and hence delay the time at which the effects of $\epsilon \neq 0$ are felt.

- The third effect however is somewhat subtle and will play an important role below in determining a bound on the mass loss due to backreaction for such a trajectory. This effect is the fact that the increase in u_*/M is *generically less* than the increase in κ^{-1}/M. In other words, for two values $\kappa_2 < \kappa_1$, although we will have $u_{*2} > u_{*1}$, we also see that $\kappa_2 u_{*2} < \kappa_1 u_{*1}$. Indeed, Fig. 4(a) shows that for the chosen values of κ, $\kappa_1 u_{*1} > 1$ whereas $\kappa_2 u_{*2} \simeq 1$. To understand the reason behind this behaviour, we note that the value of u_* for some κ and ϵ is essentially determined by the condition $e^{\alpha(u_*)} \sim A/\epsilon$ which follows from the structure of the right hand side in Eqn. (44). Also, $\alpha(u)$ is an integral of a function $h(u)$ which in turn is controlled by κ. If we reduce κ, then even though $h(u)$ takes a longer time to rise to its maximum value of 1, the integral of $h(u)$ can pick up a sizeable portion of its required budget of order $\sim M\ln(A/\epsilon)$ even for $u < \kappa_2$. This makes it plausible that although $u_{*2} > u_{*1}$, we in fact obtain $\kappa_2 u_{*2} < \kappa_1 u_{*1}$. It is difficult to give a more precise analytical argument for this observed fact.

b. A bound on mass loss

The upshot is that reducing κ also reduces the maximum magnitude of the U-dependent backreaction term. We can use this fact to construct an upper bound on the mass loss due to backreaction along this "asymptotic" trajectory. Since the V-dependent term remains subdominant until $u_{\rm ent} \sim \kappa^{-1}$, and thereafter is expected to follow a profile qualitatively similar to the U-dependent term with slightly different timescales, for order of magnitude estimates we will not consider this term separately, and instead pretend that the U-dependent term itself accounts for all the backreaction for the entire duration starting from $u = 0$. Our estimate for mass loss is a simple one : we ask for the maximum possible backreaction over the maximum possible time, and take the product. This will serve as a proxy for the integral $\int \langle T_{uu} \rangle^{\rm traj} du$ (which is the mass loss at leading order, see Sec. III).

To begin with, we notice that for any value of κ, the mass loss will be enhanced for smaller values of ϵ, since reducing ϵ will in principle allow the backreaction to be sustained at its maximum value of $\sim M^{-2}$ (although the value of κ may prevent this). This is clear from Fig. 4(a). So for any κ, the largest mass loss will be achieved when ϵ saturates its Planck scale bound (see main text, Sec. III A) of $\epsilon^2 \geq 1/2M$.

Having fixed ϵ, we see that if we choose κ appropriately so that the backreaction remains at its maximum value for times $\kappa^{-1} \lesssim u \lesssim \kappa^{-1}\ln(1/\epsilon^2)$, then the mass loss is approximated by $\Delta M \sim M^{-2}\kappa^{-1}\ln(1/\epsilon^2)$. The question now is what value of κ will maximize this quantity. Naively one might think that making κ *smaller* will help the case, since reducing κ increases the time over which the backreaction is in effect. But as we saw in the previous subsection, this is not the whole story and there is a competing effect at play : reducing κ will also mean reducing the peak magnitude of the backreaction, thus decreasing the mass loss. Clearly there is an optimum value for κ which will maximize the mass loss. Fig. 4(b) and similar figures obtained by varying κ for fixed ϵ suggest that for values of R_0 significantly larger than $2M$, this optimum value is in fact the *maximum* value $\kappa_{\rm max}$, which is also borne out by numerically integrating the U-dependent term for various values of κ (and ϵ as well).

Finally, we note that $\kappa_{\rm max}$ is a decreasing function of the starting radius R_0, which we estimated earlier as $\kappa_{\rm max} \sim M^{-1}(M/R_0)^b$ for R_0 significantly larger than $2M$, where $b > 0$ in general and $b \approx 2$ for $h(u) = \theta(u)\tanh(\kappa^2 u^2)$. If this scaling were true for all values of R_0, then a simple estimate for the largest possible value of the backreaction would come from setting $\kappa = \kappa_{\rm max}$ with $R_0/2M \gtrsim \mathcal{O}(1)$ (for an ϵ which saturates the Planck scale bound), i.e. by setting $\kappa \sim M^{-1}$ up to numerical factors. However this scaling does not hold when R_0 is close to $2M$ and $\kappa_{\rm max}$ in this case can actually be very large. Physically this is because it is possible for the trajectory to start its asymptotic phase very quickly without becoming superluminal. It might then seem that asymptotic trajectories which start close to $r = 2M$ and very quickly enter their asymptotic phase will radiate away a negligible amount of mass. This is not true since the form of the *backreaction* in the initial stages of collapse also changes in this case : notice that the U-dependent term (44) depends on $\alpha'' = h'/4M$, and for large enough κ, h approaches the Heaviside step function so that h' shows a spike close to $u = 0$ of strength $\sim M^{-1}$. The bottomline is that even for small values of R_0 for which $\kappa_{\rm max}$ can be very large, the mass loss is still of the same order of magnitude as it would be in our estimate $M^{-2}\kappa^{-1}\ln(1/\epsilon^2)$, with $\kappa \sim M^{-1}$ and $\epsilon^2 \sim M^{-1}$. The final upper bound on the mass loss (ignoring numerical factors) is given by

$$\Delta M_{\rm max} = \mathcal{O}(M^{-1}\ln(M))\,. \tag{B9}$$

3. Stress tensor VEV in initial phase of trajectory Eqn. (39)

Here we show that for type I rays which enter at times $0 < u \lesssim \kappa^{-1}\ln(1/\epsilon^2)$, the stress tensor VEV is at most of order $\mathcal{O}(1/M^2)$ and changes on a timescale $\sim M$. The argument relies on simple order of magnitude estimates using the fact that the velocity in proper time $\beta \equiv dR_s/d\tau$ in this phase satisfies $\beta = \mathcal{O}(1)$ as we show below. [β actually

starts from zero, rises in magnitude, and then falls as the asymptotic phase begins, but for simplicity we will treat it as order unity throughout this phase.] We only need to concentrate on the second term in Eqn. (27) since we already have an exact expression (44) for the first.

Writing the trajectory matching equations (6) in terms of the proper time $d\tau^2 = -ds^2|_{\rm traj}$ and noting that β is negative for a collapsing trajectory, we have the following relations

$$\frac{dV}{d\tau} = \sqrt{1+\beta^2} + \beta \;\; ; \;\; \frac{dv}{d\tau} = \frac{1}{(1-2M/R_s)}\left(\sqrt{1-2M/R_s+\beta^2}+\beta\right) , \tag{B10}$$

$$\frac{dU}{d\tau} = \sqrt{1+\beta^2} - \beta \;\; ; \;\; \frac{du}{d\tau} = \frac{1}{(1-2M/R_s)}\left(\sqrt{1-2M/R_s+\beta^2}-\beta\right) . \tag{B11}$$

Using these, one can write down expressions for U' and $\dot{V}$ in terms of β and R_s. Further, in the initial phase of the trajectory we have $U' = \epsilon + Ae^{-\alpha(u)} = \mathcal{O}(1)$, $(1-2M/R_s) = \mathcal{O}(1)$ and hence $dv/du = \mathcal{O}(1)$. This implies that β and hence $\dot{V}$ are of order unity.

The second term in Eqn. (27) can therefore be estimated as $\sim \tau_V^{-2}$ where τ_V is the timescale controlling $\dot{V}$. This timescale is set entirely by the function U' due to the way we have chosen to define the trajectory. We can therefore estimate τ_V^{-1} as

$$\tau_V^{-1} \sim \partial_u \ln U' \sim \alpha' = h(u)/4M , \tag{B12}$$

other factors being of order unity. Due to the required asymptotic behaviour of $h(u)$, we can approximately take $h(u) \sim 0$ for $u \lesssim \kappa^{-1}$ and $h(u) \sim 1$ for $\kappa^{-1} \lesssim u \lesssim \kappa^{-1}\ln(1/\epsilon^2)$. This is not too bad an approximation since, for example, if $h(u) = \tanh(\kappa^2 u^2)$, we will have $h \sim \kappa^2 u^2$ for $u \ll \kappa^{-1}$ and $h \sim 1 + \mathcal{O}(e^{-2\kappa^2 u^2})$ for $u \gg \kappa^{-1}$. This estimate shows that the second term $(U'/\dot{V})^2 \dot{V}^{1/2}\partial_v^2 \dot{V}^{-1/2}$ is expected to be significant during the interval $\kappa^{-1} \lesssim u_{\rm ent} \lesssim \kappa^{-1}\ln(1/\epsilon^2)$ in which its value is at most $\mathcal{O}(1/M^2)$ and changes on a timescale $\sim M$.

4. Stress tensor VEV in asymptotic phase of trajectory Eqn. (39)

Here we show that the second term in Eqn. (27) exponentially decays for the type II rays which enter and exit the shell in the asymptotic phase $u \gg \kappa^{-1}\ln(1/\epsilon^2)$ of the trajectory Eqn. (39). Using the late time solution (B7), we can parametrically obtain $v_{\rm ent}(V)$ by finding $v(u)|_{\mathcal{P}_{\rm ent}}$ and $V(u)|_{\mathcal{P}_{\rm ent}}$ as follows : From Eqns. (6) we get

$$\frac{dv}{du} = 1 + \frac{2R_s'}{1-2M/R_s} = 1 + \mathcal{O}\left(\frac{\tilde{A}}{\epsilon^2}e^{-\mathcal{K}u/4M}\right) \;\; ; \;\; \dot{V} = \frac{1}{U'}(1-2M/R_s) = \epsilon\left(1+\mathcal{O}\left(\frac{\tilde{A}}{\epsilon^2}e^{-\mathcal{K}u/4M}\right)\right) , \tag{B13}$$

where $\mathcal{K} = \mathcal{O}(1)$ is determined from the solution (B7) and $\tilde{A} = \mathcal{O}(e^{1/\kappa M})$ was defined in Appendix B 1 above. This leads to

$$v_{\rm ent}(V) = \frac{V}{\epsilon} + \mathcal{O}\left(\frac{M\tilde{A}}{\epsilon^2}e^{-\mathcal{K}V/4M\epsilon}\right) . \tag{B14}$$

Further, since the exit point is also in the asymptotic phase, we have

$$U = \epsilon\left(u + \mathcal{O}\left(\frac{M\tilde{A}}{\epsilon}e^{-u/4M}\right)\right) , \tag{B15}$$

which leads to

$$v_{\rm ent}(U(u)) = u + \mathcal{O}\left(\frac{M\tilde{A}}{\epsilon^2}e^{-\mathcal{K}u/4M}\right) . \tag{B16}$$

Using the expression for $\dot{V}$ from Eqn. (B13) above and replacing u (which is actually $u_{\rm ent}$) by v at the leading order, one can easily show that

$$\dot{V}^{1/2}\partial_v^2 \dot{V}^{-1/2} = \mathcal{O}\left(\frac{\tilde{A}}{\epsilon^2 M^2}e^{-\mathcal{K}v/4M}\right) , \tag{B17}$$

and further, evaluation at $v = v_{\rm ent}(U(u))$ simply replaces v by u (which is now $u_{\rm ex}$). Also, at the leading order we have $U'^2/\dot{V}^2 \approx 1$, and hence the second term in Eqn. (27) becomes

$$\frac{U'^2}{\dot{V}^2}\,\dot{V}^{1/2}\partial_v^2\dot{V}^{-1/2}\Big|_{v=v_{\rm ent}(U(u))} = \mathcal{O}\left(\frac{\tilde{A}}{\epsilon^2 M^2}e^{-\mathcal{K}u/4M}\right)\,, \tag{B18}$$

with the approximation valid for $u \gg \kappa^{-1}\ln(1/\epsilon^2)$.

5. The effect of backreaction on the background geometry

In this section we sketch the arguments presented by Brout et al. which show that the late time backreaction is determined by the time dependent mass $M(u)$ of the collapsing object. The reader is referred to Sec 3.4 of Ref. [3] for further details. The argument proceeds as follows : The spherically symmetric exterior metric is written in the form

$$ds^2_{\rm ext} = -e^{2\psi}\left(1-\frac{2m(v,r)}{r}\right)dv^2 + 2e^{\psi}dvdr + r^2d\Omega^2\,, \tag{B19}$$

in terms of which the Einstein equations become

$$\frac{\partial m}{\partial v} = T^r_{\ v} \;\;;\;\; \frac{\partial m}{\partial r} = -T^v_{\ v} \;\;;\;\; \frac{\partial \psi}{\partial r} = T^r_{\ r}/r\,, \tag{B20}$$

where the right hand sides contain the *two-dimensional* stress tensor components.

At large distances, the zeroth order calculation shows that only $T_{uu} = L_H$ survives. (We have already computed T_{vv} which falls like $\sim (M/r)^3$, and the energy-momentum conservation equation can be used to show that so does T_{uv}.) Hence the metric must correspond to the outgoing Vaidya solution

$$ds^2_{{\rm ext,large}\,r} = -\left(1-\frac{2M(u)}{r}\right)du^2 - 2dudr + r^2d\Omega^2 \;\;;\;\; \frac{dM(u)}{du} = -T_{uu}\,, \tag{B21}$$

where we assume that $L_H = \mathcal{O}(1/M^2)$ and varies slowly. Matching (B19) and (B21) at some large radius (say $r > \mathcal{O}(6M)$) gives the transformation between the (v,r) and (u,r) coordinates as

$$e^{\psi}dv = du + \frac{2dr}{1-2M(u)/r} \;\;;\;\; m(v,r) = M(u)\,. \tag{B22}$$

Using this to calculate the right hand sides of the Einstein equations (B20) in terms of L_H, one finds that at large r all three stress tensor components $T^r_{\ v}$, $T^v_{\ v}$ and T_{rr}, are of order $\mathcal{O}(L_H)$ times some quantity of order unity.

Near the apparent horizon $r_{\rm ah}(v) = 2m(v, r_{\rm ah}(v))$, one uses the fact that r and v both behave like inertial light-like coordinates (see Eqn. (B19)), to argue that $T^v_{\ v}$ and T_{rr} must be of order $\mathcal{O}(L_H)$ here as well. Integrating the conservation equation $T^r_{\ v,r} + T^v_{\ v,v} = 0$ from $r_{\rm ah}$ to $r > \mathcal{O}(6M)$ then gives $T^r_{\ v} = \mathcal{O}(L_H)$ near the horizon also. Specifically, one finds near the horizon (from Eqn. 3.86 of Ref. [3])

$$\frac{\partial m}{\partial v} = -L_He^{\psi} + \mathcal{O}(ML_{H,v}) \;\;;\;\; \frac{\partial m}{\partial r} = \mathcal{O}(L_H) \;\;;\;\; \frac{\partial \psi}{\partial r} = \mathcal{O}(L_H/r)\,, \tag{B23}$$

which shows that the metric function m varies slowly and also that ψ can be safely set to zero in the calculation.

The next step is to obtain the equation for outgoing null geodesics (ONGs). This is somewhat involved due to the presence of the apparent horizon where the ONGs change the sign of their "velocity" dr/dv. The analysis proceeds by first noting that as long as L_H is small and slowly varying, the location of the *event horizon* $r = r_{\rm eh}(v)$ is not very far removed from that of the apparent horizon. To see this, we write $r_{\rm eh}(v) = r_{\rm ah}(v) + \Delta(v)$. Differentiating with respect to v and using $2dr_{\rm eh}/dv = 1 - 2m(v, r_{\rm eh})/r_{\rm eh}$ (since the event horizon by definition is the last ONG to reach $\mathcal{I}^+$), we find at leading order in L_H

$$\Delta = r_{\rm ah}r_{{\rm ah},v} \sim -1/M \Rightarrow r_{\rm eh} = r_{\rm ah}(1+\mathcal{O}(L_H))\,, \tag{B24}$$

since $dr_{\rm ah}/dv = \mathcal{O}(L_H)$. A useful result which follows from this is $2m(v, r_{\rm eh}) = r_{\rm eh}(1+\mathcal{O}(L_H))$,

$$2m(v,r_{\rm eh}) = 2m(v, r_{\rm ah}+\Delta) = 2m(v,r_{\rm ah}) + \Delta\mathcal{O}(L_H) + \ldots = r_{\rm ah}(1+\mathcal{O}(L_H^2)) = r_{\rm eh}(1+\mathcal{O}(L_H))\,. \tag{B25}$$

The ONGs are easier to analyse in terms of the coordinate $x = r - r_{\rm eh}(v)$, in terms of which the metric becomes

$$ds^2_{\rm ext} = -dv^2 \frac{2m(v, r_{\rm eh} + x)x}{r_{\rm eh}(r_{\rm eh} + x)}(1 + \mathcal{O}(L_H)) + 2dvdx + r^2 d\Omega^2 \,. \tag{B26}$$

The coefficient g_{vv} vanishes on the *event horizon* ($x = 0$) in these coordinates rather than the apparent horizon, and the ONGs near the horizon resemble those in the unperturbed Schwarzschild geometry, namely we have for $x \ll r_{\rm eh}$,

$$\tilde{v} - 2\ln x = f(u) \quad ; \quad \tilde{v} = \int^v \frac{dv}{r^2_{\rm eh}} 2m(v, r_{\rm eh})(1 + \mathcal{O}(L_H))\,, \tag{B27}$$

along an ONG, and the following ansatz describes ONGs at arbitrary distances x,

$$\tilde{v} - 2\frac{x}{r_{\rm eh}(v)} - 2\ln x + \delta = \int^u \frac{du'}{\tilde{m}(u')} + D\,. \tag{B28}$$

Here D is a constant of integration, and $\delta = \mathcal{O}(L_H(Mx + x^2)/M)$ which follows from integrating the ONG equation $u =$const using the form of the metric (B26). By requiring that at large distances the coordinate u be identical to the one appearing in the outgoing Vaidya metric (B21), one finds $\tilde{m}(u) = M(u)(1+\mathcal{O}(L_H))$ using the transformation equation (B22).

This coordinate transformation is then used to calculate the quantity $\langle T_{uu} \rangle^{\rm ren}$ in the same way as was done in the case without backreaction, and to leading order one finds that the flux on the horizon vanishes quadratically and the flux reaching $\mathcal{I}^+$ is $(\pi/12)T_H^2(u)$ where

$$T_H(u) \simeq \frac{1}{8\pi\tilde{m}(u)} = \frac{1}{8\pi M(u)}(1 + \mathcal{O}(L_H))\,, \tag{B29}$$

which completes the argument (see Eqns. 3.91-3.95 of Ref. [3] for details).

6. Delay in horizon formation for a null trajectory

In this section we analyse the effect of backreaction on the horizon formation time for a null in-falling trajectory. Our classical trajectory in the absence of backreaction has two phases :

1. $u < 0$ **:** $R_s(u) = R_0 =$const.
2. $u > 0$ **:** $v =$const., $V =$const.

The horizon is formed at finite (U, V). The trajectory solution is

$$\textbf{Phase (1):}\; U(u) = \left(1 - \frac{2M}{R_0}\right)^{1/2} u \,-\, 2R_0 + 4M \quad ;$$
$$V(v) = \left(1 - \frac{2M}{R_0}\right)^{1/2} v \,-\, 2x(R_0)\left(1 - \frac{2M}{R_0}\right)^{1/2} + 4M \quad ; \quad v = u + 2x(R_0)\,, \tag{B30}$$
$$\textbf{Phase (2):}\; V = 4M \quad ; \quad v = 2x(R_0) \quad ;$$
$$U = -2R_s + 4M \quad ; \quad u = -2x(R_s) + 2x(R_0)\,, \tag{B31}$$

where $x(r) \equiv r + 2M\ln(r/2M - 1)$ is the "tortoise" function and constants are chosen so that the horizon is formed at $U = 0$ (with $u \to \infty$). The Penrose diagram for this spacetime is shown in Fig. 5. Causality implies that backreaction cannot have any effect on the spacetime region $U < 4M - 2R_0$, i.e. while the trajectory is in phase (1). In phase (2) the trajectory will be modified and will no longer in general have $V = 4M =$const. Say the modified trajectory (parametrized by U) is given by $(V = \tilde{V}(U), r = \tilde{R}_s(U))$ in interior coordinates and $(v = \tilde{v}(U), r = \tilde{R}_s(U))$ in exterior coordinates. The interior geometry is still Minkowski while the exterior is the Brout et al. approximation

$$ds^2_{\rm ext} \approx -\left(1 - \frac{2m(v, r)}{r}\right) dv^2 + 2dvdr + r^2 d\Omega^2\,, \tag{B32}$$

where the mass function $m(v, r)$ is slowly varying in the entire exterior. All this refers to phase (2) of the trajectory. The matching therefore leads to

$$d\tilde{V} - dU = 2d\tilde{R}_s \quad ; \quad -dU^2 - 2dUd\tilde{R}_s = -\left(1 - \frac{2m(\tilde{v}, \tilde{R}_s)}{\tilde{R}_s}\right) d\tilde{v}^2 + 2d\tilde{v}d\tilde{R}_s\,, \tag{B33}$$

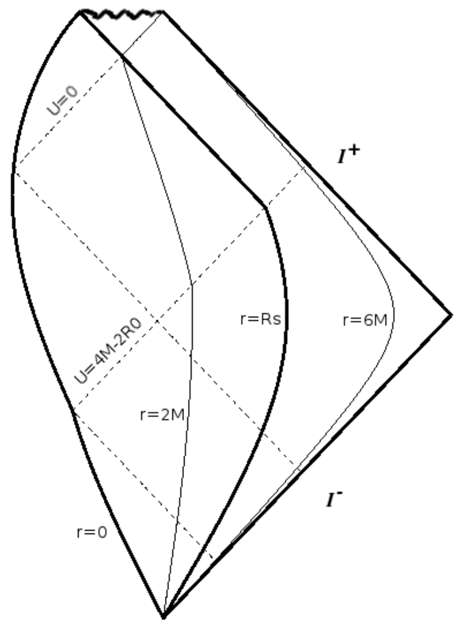


FIG. 5: Penrose diagram (in the absence of backreaction) for a shell which stays at $r = R_0 = 3.5M$ until $U = 4M - 2R_0$ and then collapses along $v =$const., forming a horizon at $U = 0$. The shell trajectory and the $r = 2M$ surface are labelled. The two null rays marking the onset of phase (2) and horizon formation, are also shown.

and we will assume that the functional form of the trajectory is shifted by terms of order $\mathcal{O}(L_H)$ so that

$$\tilde{R}_s(U) = R_s(U)(1 + \mathcal{O}(L_H)) \quad ; \quad R_s(U) = -\frac{1}{2}(U - 4M_0)\,, \tag{B34}$$

with M_0 denoting the unperturbed mass of the shell. For the backreaction L_H, *at the leading order* we will take this to be given by the exact calculation of Sec. II B (see Eqn. (37)), so that

$$L_H = \langle\, T_{uu}\,\rangle^{\rm traj}(R_s) = \frac{1}{48\pi}\frac{M_0}{R_s^3}\left(2 - \frac{3M_0}{R_s}\right)\,. \tag{B35}$$

Although the calculation of Sec. II B was for a trajectory which does not form a horizon, the derivation can be easily modified to show that the stress tensor VEV (in the absence of backreaction) in the present case, will have the form (B35) at least for $R_s > 2M_0$. Using this form for L_H is self-consistent so long as $L_H \ll 1$, which is the case for all $R_s \geq 2M_0$. We then have

$$1 + 2\frac{d\tilde{R}_s}{dU} = \mathcal{O}(L_H)\,, \tag{B36}$$

which gives us, at the leading order,

$$\frac{d\tilde{V}}{dU} = \mathcal{O}(L_H) \quad ; \quad \frac{d\tilde{v}}{dU} = \mathcal{O}(L_H)\,. \tag{B37}$$

Now assume that the horizon forms at $U = \delta U$ instead of $U = 0$, so that

$$\tilde{R}_s(\delta U) = r_{\rm eh}(\tilde{v}(\delta U)) \quad ; \quad \tilde{v}(\delta U) \equiv v_0 + \delta v\,, \tag{B38}$$

where $v_0 = 2x(R_0)$ is the unperturbed value at which the horizon is formed. We would like to estimate δU and δv. Using Eqn. (B37) we have

$$\tilde{v}(\delta U) = v_0 + \int_{-2R_0+4M_0}^{\delta U} \mathcal{O}(L_H)dU\,, \tag{B39}$$

where the lower limit of integration is set by requiring that the trajectories in the presence and absence of backreaction be identical at the beginning of phase (2) which is at $U = -2R_0 + 4M_0$. At this level of approximation, using the

expression (B35) for L_H with $R_s(U)$ given by Eqn. (B34), we find

$$\int_{-2R_0+4M_0}^{\delta U} L_H dU = \frac{M_0}{6\pi}\frac{(2M_0-U)}{(4M_0-U)^3}\bigg|_{-2R_0+4M_0}^{\delta U}\,, \tag{B40}$$

and hence, ignoring numerical factors,

$$\frac{\delta v}{M} = \mathcal{O}(M^{-2})\left(1+\mathcal{O}(\delta U/M)+\mathcal{O}((M/R_0)^2)\right)\,. \tag{B41}$$

We can further assume $r_{\rm eh}(\tilde{v}(\delta U)) = 2M_0(1+\mathcal{O}(M^{-2}))$ since $r_{\rm eh}(v)$ is a slowly varying function whose unperturbed form is the constant $2M_0$, and since $L_H = \mathcal{O}(M^{-2})$ around horizon formation. Then Eqn. (B38) gives us

$$\delta U/M = \mathcal{O}(M^{-2})\,, \tag{B42}$$

so that $|\delta U|/M \ll 1$ as expected. The signs of δv and δU are not determined by this analysis. As in the timelike case, the backreaction in this case cannot halt the horizon formation, but can only delay it by a time $\mathcal{O}(1/M)$.

---

[1] S Hawking, *Nature* **248**, 30 (1974); *Commun. Math. Phys.* **43**, 199 (1975).
[2] M Visser, *Int. J. Mod. Phys.* **D12**, 649 (2003) [arXiv:hep-th/0106111].
[3] R Brout, S Massar, R Parentani and Ph Spindel, *Phys. Rept.* **260**, 329 (1995) [arXiv:0710.4345].
[4] N D Birrell and P C W Davies, *Quantum Fields in Curved Space*, Cambridge Univ. Press, Cambridge (1984).
[5] D N Page, *Phys. Rev.* **D13**, 198 (1976).
[6] N G Sanchez, *Phys. Rev.* **D18**, 1030 (1978).
[7] K W Howard, *Phys. Rev.* **D30**, 2532 (1984).
[8] P C W Davies, S A Fulling and W G Unruh, *Phys. Rev.* **D13**, 2720 (1976).
[9] J M Bardeen, *Phys. Rev. Lett.* **46**, 382 (1981).
[10] J W York, Jr., *Phys. Rev.* **D28**, 2929 (1983).
[11] W G Unruh, *Phys. Rev.* **D14**, 870 (1976).
[12] S Massar, *Phys. Rev.* **D52**, 5857 (1995) [arXiv:gr-qc/9411039].
[13] R Brout, et al., *Phys. Rev.* **D52**, 1119 (1995) [arXiv:hep-th/9311019].
[14] T Clifton, *Class. Quant. Grav.* **25**, 175022 (2008) [arXiv:0804.2635].
[15] B A Brown and J Lindesay, *Class. Quant. Grav.* **25**, 105026 (2008) [arXiv:0710.2032].
[16] A B Nielsen and D-h Yeom, arXiv:0804.4435 (2008).
[17] T Vachaspati, D Stojkovic and L M Krauss, *Phys. Rev.* **D76**, 024005 (2007) [arXiv:gr-qc/0609024];
T Vachaspati and D Stojkovic, *Phys. Lett.* **B663**, 107 (2008) [arXiv:gr-qc/0701096];
T Vachaspati, arXiv:0711.0006 (2007).
[18] C Barcelo, S Liberati, S Sonego and M Visser, *Phys. Rev.* **D77**, 044032 (2008) [arXiv:0712.1130].
[19] M Visser, arXiv:0901.4365 (2009), contributed to Workshop on Black Holes in General Relativity and String Theory, Veli Losinj, Croatia, 24-30 Aug 2008.
[20] G L Alberghi, R Casadio, G P Vacca and G Venturi, *Phys. Rev.* **D64**, 104012 (2001) [arXiv:gr-qc/0102014].
[21] T P Singh and T Padmanabhan, *Ann. Phys.* **196**, 296 (1989);
T Padmanabhan, *Class. Quant. Grav.* **6**, 533 (1989);
J B Hartle, *Phys. Rev.* **D38**, 2985 (1988).
[22] For completeness, we also note that had we done the analysis with $\epsilon = 0$ exactly, we would find that as $u \to \infty$, $v_{\rm ent}(V)$ linearises around the "last escaper" ray (see Appendix A 1), and the contribution from the second term is suppressed like $\sim e^{-u/2M}$. In practice, since we can expect $u_{\rm ex}(u_{\rm ent}=0) > 2M$, the second term will remain highly suppressed at all times. The Hawking flux therefore arises entirely from the first term in Eqn. (27), in the standard case.]
[23] T Padmanabhan, *Phys. Rept.* **406**, 49 (2005) [arXiv:gr-qc/0311036].